Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 42 (2007) 1365–1369
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

An atom counting strategy towards analyzing the biological activity of sex hormones

D.R. Roy [a], N. Pal [a], A. Mitra [b], P. Bultinck [c], R. Parthasarathi [d], V. Subramanian [d], P.K. Chattaraj [a,*]

[a] *Department of Chemistry, Indian Institute of Technology, Kharagpur 721302, India*
[b] *B. C. Roy Technology Hospital, Indian Institute of Technology, Kharagpur 721302, India*
[c] *Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281, B-9000 Gent, Belgium*
[d] *Chemical Laboratory, Central Leather Research Institute, Adyar, Chennai 600020, India*

Received 27 November 2006; received in revised form 25 January 2007; accepted 29 January 2007
Available online 24 February 2007

Abstract

A simple and effective molecular descriptor, viz., the number of atoms in a molecule (N_A) is made use of in the development of the quantitative structure–activity relationship (QSAR). A series of testosterone derivatives with various biological activities and estrogen derivatives with the activities in terms of relative binding affinity (RBA) are considered to find out the potential of N_A in predicting the activities of those molecules. It is heartening to note that N_A along with the electrophilicity index (ω) is capable of explaining the biological activities of the male and female hormones.

Keywords: Atom counting; QSAR; Electrophilicity; Testosterone; Estrogen

1. Introduction

Hormones are the chemical messengers carrying information from cell to cell in all multicellular organisms including plants (phytohormone). The most important human hormones may be classified into four groups [1–2], viz. (i) amine (derived) hormones i.e. tryptophan derivatives (e.g. melatonin or *N*-acetyl-5-methoxytryptamine) and tyrosine derivatives (e.g. thyroxine), (ii) peptide hormones (e.g. calcitonin, insulin), (iii) lipid and phospholipid hormones (eicosanoids) (e.g. prostaglandins, thromboxane) and (iv) sterol and steroid hormones. Vitamin D derivatives and calcitriol are the examples of sterol hormones. Steroidal hormones may be classified into three groups, viz., glucocorticoids (e.g. cortisol), mineralocorticoids (e.g. aldosterone) and sex steroids. Sex steroids include androgens (e.g. testosterone, androstenedione etc.), estrogens (e.g. estradiol) and progestagens (e.g. progesterone, progestins). Sex steroid hormones are the most responsible for control and development of male and female characteristics. Testosterone, the major androgen, is important for the development of the male reproductive tract fertility and secondary male characteristics [1–3]. It is the principal sex hormone of male and an anabolic steroid, secreted in the testes of males and the ovaries of females although small amounts are secreted by the adrenal glands. Estradiol, the major circulating estrogen, is produced in the ovaries of females. Estradiol is important in the female reproductive tract, playing roles in breast and uterine development, fertility and in the maintenance of pregnancy and also in bone and other tissues. These hormones act on a target tissue through binding with a part of cell called their receptors [1–3].

The general mechanism of the action of these hormones may be understood as follows. As all the ligands are lipophilic compounds, they enter into the cell mainly facilitated by

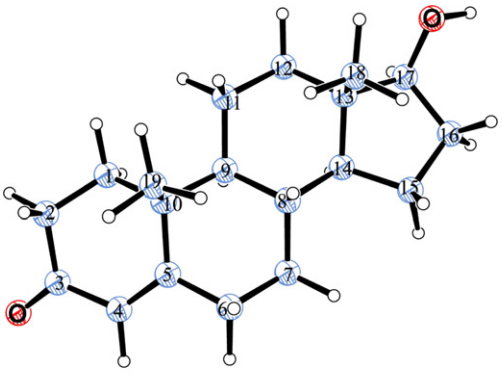
* Corresponding author.
E-mail address: pkc@chem.iitkgp.ernet.in (P.K. Chattaraj).

0223-5234/$ - see front matter © 2007 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2007.01.028

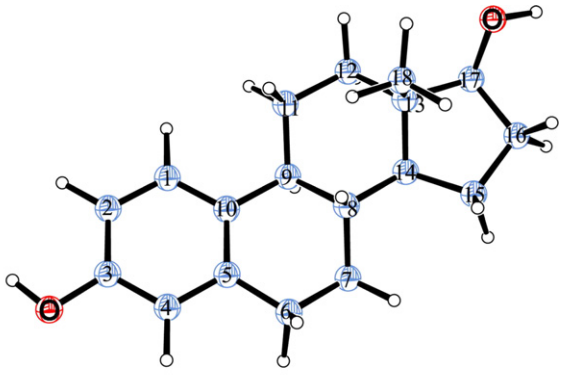
passive diffusion. They bind to their cognate intracellular receptors located either in the cytoplasm or nucleus [3]. The combined unit of the hormone and the receptor moves across the nuclear membrane into the nucleus of the cell and binds to the DNA, thereby effectively amplifying or suppressing the action of certain genes, which in turn affects the protein synthesis [1,2]. Although steroid receptor family members like the estrogens, androgens and progesterones are important for normal physiological processes, there are a number of instances in which it is desirable to block the actions of selected steroid receptors. These include breast cancer (estrogen receptor) and prostrate cancer (androgen receptor). Thus, although natural antagonists of steroid receptor action have not been identified, much effort has been devoted in identifying compounds that will antagonize hormone's action. These compounds compete with the natural ligand for binding to the hormone-binding domain of the receptor [3].

The interaction between the hormones and the respective receptors may be scaled as the various biological activities, e.g. relative binding affinity (RBA), androgenic potency (AP), therapeutic index (TI) etc. The quantitative structure–activity relationship (QSAR) is a bridge between the structure of a molecule and its desired biological activity through a mathematical representation [4–8]. Once a correlation is established, the biological activities of any number of molecules/drugs with the similar molecular structure can be predicted. Thus it is possible to save resources, viz., time and money through the QSAR study. In certain cases even the unavailability of molecules causes impediments towards understanding the actual biological activity. Numerous descriptors based on topology, thermodynamics, quantum chemistry, electronic energy etc. have been used in the development of QSAR [4–8]. The density functional theory (DFT) [9,10] based quantum chemical descriptors have gained a lot of attention recently due to their reliability and versatility in the prediction of biological activity/toxicity [9–11]. Electrophilicity index (ω) [12,13] and philicity [14] are found to be quite successful in describing the toxicity of polychlorinated biphenyls (PCBs) and benzidine [15], polychlorinated dibenzofurans (PCDFs), polychlorinated dibenzo-*p*-dioxins (PCDDs) [16] and chlorophenols (CP) [17]. Also, acute toxicity of diverse classes of aliphatic [18] and aromatic [19] molecules towards the ciliated freshwater protozoa *Tetrahymena pyriformis* is predicted successfully with the use of global and local electrophilicities. Biological activities of a series of testosterone and estradiol derivatives [20] are also explained successfully by electrophilicity index (ω).

The purpose of the present work is to find out the biological activity of the different sets of testosterone and estrogen derivatives using a very simple descriptor, viz., the number of atoms in a molecule (N_A). Testosterone derivatives with various biological activities in terms of androgenic potency (AP) [21], therapeutic index (TI) [22] and testosterone–estrogen binding globulin (TeBG) [23] whereas estrogen derivatives with the relative binding affinity (RBA) [24] are selected as the model system to find out the usefulness of the number of atoms in a molecule (N_A) as a potent descriptor in explaining their biological activities.



Structure of Testosterone



Structure of Estrogen

Section 2 provides the theoretical background followed by Section 3 presenting the computational details. Results and discussion are provided in Section 4 and finally Section 5 contains some concluding remarks.

2. Theoretical background

The number of atoms in a molecule (N_A) has been considered [25,26] to be a valid descriptor in different occasions. In the present work we explore the possibility of using N_A as the descriptor of biological activities when a molecule interacts with the corresponding receptor. We choose the activities of various sex hormones for this purpose. Since global and local electrophilicity indices are known [20] to be useful descriptors of biological activities of testosterones and estrogens, in order to improve the situation with N_A as the descriptor, electrophilicity index (ω) is also used as an additional descriptor in the study. Parr et al. [12] defined electrophilicity index as

a measure of energy lowering due to maximal electron flow between a donor and an acceptor as follows:

$$\omega = \frac{\mu^2}{2\eta} \quad (1)$$

where we write, using a finite difference approach, $\mu \approx -(I+A)/2$ and $\eta \approx (I-A)/2$ which are the electronic chemical potential and the chemical hardness of the ground state of atoms or molecules, respectively, approximated in terms of the vertical ionization potential (I) and electron affinity (A). Using the Koopmans' theorem for closed-shell molecules I and A can be expressed in terms of the highest occupied molecular orbital energy (ϵ_{HOMO}) and the lowest unoccupied molecular orbital energy (ϵ_{LUMO}) as follows [9]:

$$I \approx -\epsilon_{HOMO};\ A \approx -\epsilon_{LUMO} \quad (2)$$

The usefulness of N_A in the prediction of biological activity is tested for the three different sets of testosterones and one set of estrogen derivatives in terms of various biological activities.

3. Computational details

Three different groups of testosterone derivatives in terms of various biological activities e.g. androgenic potency (AP) [21], therapeutic index (TI) [22] and testosterone–estrogen binding globulin affinity (TeBG) [23] and estrogen derivatives with the relative binding affinity (RBA) [24] are chosen for the study. All the molecules are minimized at the Austin model 1 (AM1) level of theory followed by a single point calculation at the B3LYP/6-31G* level of theory with AM1 geometry using the Gaussian 03W [27]. Electrophilicity index (ω) is calculated using the Eqs. (1) and (2). A QSAR study is performed using the least square error estimation method. The biological activities (AP, TI, TeBG and RBA) are taken as the dependent variables whereas the number of non-hydrogenic atoms (N_{NH}) and electrophilicity index (ω) are used as independent variables. One parameter regression with N_{NH} is performed to find out the biological activity of the various sets of testosterone and estrogen derivatives. A two parameter model is also developed with the combination of N_{NH} and ω.

4. Results and discussion

A total of 31 testosterone derivatives with various biological activities, e.g. androgenic potency (AP) [21], therapeutic index (TI) [22] and testosterone–estrogen binding globulin affinity (TeBG) [23] and 12 estrogen derivatives with the relative binding affinity (RBA) [24] are used to find out the potential of the number of non-hydrogenic atoms (N_{NH}) as a valid descriptor in the QSAR parlance. One and two parameter regression analyses are performed using N_{NH} and ω to find out the biological activities of the selected testosterone and estrogen derivatives.

4.1. Testosterone derivatives

Tables 1–3 present the different sets of testosterone derivatives with their experimental and calculated biological activities in terms of AP, TI and TeBG. The corresponding one and two parameter regression models for each group are provided in Table 4. Electrophilicity values are reported up to four decimal places in all the tables. It may be noted that N_{NH} can explain the activities of all the three groups in an acceptable manner. The potential of ω [12,13] is then tested by combining it with N_{NH} to form a two parameter QSAR model. The combination of ω and N_{NH} drastically improves the correlation implying the important role played by ω as well [20]. Therefore two parameter model with ω and N_{NH} may be considered as the best model for the testosterone derivatives to explain their biological activities. Both ω and N_{NH} are important descriptors in analyzing the activities of the testosterone derivatives.

4.2. Estrogen derivatives

Table 5 provides the estrogen derivatives with the experimental and calculated biological activities in terms of relative binding affinity (RBA). The one parameter (N_{NH}) and two parameter (N_{NH}, ω) regression models, with the associated correlation coefficient (R), standard deviation (SD) and number of observations (N), to find the RBA values are as follows:

$$\text{Calc. RBA} = -0.1417 \times N_{NH} + 4.8741 \quad R = 0.82;\ \text{SD} = 0.32;\ N = 12 \quad (3)$$

Table 1
Electrophilicity index (ω) of testosterone derivatives with their experimental and calculated biological activities in terms of androgenic potency (AP) [21]

No.	Compounds	N_{NH}	Electrophilicity (ω) (eV)	Exptl AP	Calc AP	
					N_{NH}	N_{NH}, ω
1	7β-Hydroxy,17α-methyl-5α-androst-1-en-3-one	22	2.9591	25	38.08	11.77
2	17β-Hydroxy,17α-ethyl-5α-androst-1-en-3-one	23	2.9499	2	18.38	9.64
3	17β-Hydroxy,2,17α-dimethyl-5α-androst-1-en-3-one	23	2.8859	25	18.38	12.08
4	17β-Hydroxy,2-methyl,17α-ethyl-5α-androst-1-en-3-one	24	2.8777	1	−1.33	9.92
5	5α-Androst-1-en-3β,17β-diol	21	1.3304	50	57.79	76.19
6	17α-Methyl-5α-androst-1-ene-3β,17β-diol	22	1.3307	100	38.08	73.71
7	17β-Hydroxy,6β-methyl-5α -androst-1-en-3-one	22	2.9337	10	38.08	12.73
8	17β-Hydroxy,6β,17α-dimethyl-5α-androst-1-en-3-one	23	2.9327	10	18.38	10.30
9	17β-Hydroxy,6β-methyl,17α-ethyl-5α-androst-1-en-3-one	24	2.9236	1.5	−1.33	8.17

Table 2
Electrophilicity index (ω) of testosterone derivatives with their experimental and calculated biological activities in terms of therapeutic index (TI) [22]

No.	Compounds	N_{NH}	Electrophilicity (ω) (eV)	Exptl TI	Calc TI	
					N_{NH}	N_{NH}, ω
1	Methylandrostenediol	21	1.0897	0.28	0.41	0.28
2	19-Nor-17a-ethyltestosterone	21	2.6932	0.55	0.41	0.54
3	4-Fluro testosterone acetate	25	3.2184	0.58	0.59	0.61
4	4-Chloro testosterone acetate	25	3.3630	0.68	0.59	0.64
5	4-Chloro testosterone propionate	26	3.2170	0.56	0.63	0.61
6	4-Hydroxy testosterone acetate	25	3.0740	0.52	0.59	0.59
7	4-Hydroxy-19-nor-testosterone acetate	25	3.0493	0.55	0.59	0.59
8	4-Chloro-17a-methyl-19-nor-testosterone	26	3.1643	0.73	0.63	0.60

$$\text{Calc. RBA} = -0.1270 \times N_{NH} - 1.5142 \times \omega + 6.6303$$
$$R = 0.84;\ \text{SD} = 0.30;\ N = 12 \qquad (4)$$

It may be noted that N_{NH} alone is capable of providing a satisfactory explanation of the biological activity of the selected estrogen derivatives. The combined model with N_{NH} and ω (Eq. (4)) shows some better prediction. Fig. 1a and b presents the correlation and the histogram plots between the experimental and the calculated biological activities (Tables 1–5) for the complete set of 43 testosterone and estrogen derivatives. The higher coefficient of correlation (R) and the corresponding histogram representation show how beautifully the combination of N_{NH} and ω can describe the biological activities of those male and female sex hormones.

Biological activities of the testosterone and estrogen derivatives are studied in the past with the electronegativity, hardness or electrophilicity [20,28]. Although electronegativity fails to correlate with the biological activity in a meaningful way and a reasonably good correlation with hardness has been rationalized in the light of the maximum hardness principle [28],

Table 3
Electrophilicity index (ω) of testosterone derivatives with their experimental and calculated biological activities in terms of TeBG affinity (TeBG) [23]

No.	Compounds	N_{NH}	Electrophilicity (ω) (eV)	Exptl TeBG	Calc TeBG	
					N_{NH}	N_{NH}, ω
1	Androstanediol	21	0.7051	−9.11	−8.10	−9.09
2	Androstenediol	21	1.0920	−9.17	−8.10	−8.71
3	Androsterone	21	1.9028	−7.14	−8.10	−7.92
4	Corticosterone	25	2.6536	−6.34	−6.91	−6.89
5	Cortisol	26	2.9378	−6.20	−6.62	−6.53
6	Cortisone	29	2.9863	−6.41	−5.72	−6.26
7	Dehydroepiandrosterone	21	1.9679	−7.81	−8.10	−7.85
8	Deoxycorticosterone	24	2.7935	−7.38	−7.21	−6.82
9	Deoxycortisol	25	2.8711	−7.20	−6.91	−6.67
10	17β-Estradiol	20	1.3581	−8.83	−8.40	−8.53
11	Estrone	20	1.8445	−8.17	−8.40	−8.05
12	Pregnenolone	23	1.9186	−7.14	−7.51	−7.75
13	Progesterone	23	2.7423	−6.94	−7.51	−6.95
14	17-Hydroxy progesterone	24	2.8140	−6.99	−7.21	−6.80

Table 4
Regression models and the associated correlation coefficient (R), standard deviation (SD) and number of observations (N) for the different groups of testosterone derivatives with the various biological activities, viz., androgenic potency (AP), therapeutic index (TI) and TeBG affinity (TeBG)

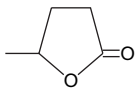
Regression models	R	SD	N
$\text{AP} = -19.708 \times N_{NH} - 471.667$	0.61	27.38	9
$\text{AP} = -2.474 \times N_{NH} - 38.036 \times \omega + 178.748$	0.88	16.52	9
$\text{TI} = 0.044 \times N_{NH} - 0.519$	0.68	0.11	8
$\text{TI} = -1.514 \times 10^{-3} \times N_{NH} + 0.161 \times \omega + 1.329 \times 10^{-1}$	0.88	0.07	8
$\text{TeBG} = 0.297 \times N_{NH} - 14.348$	0.78	0.78	14
$\text{TeBG} = 7.388 \times 10^{-2} \times N_{NH} + 0.981 \times \omega - 11.335$	0.90	0.44	14

electrophilicity plays a good role in predicting the biological activities of the male and female hormones [20]. The potential of the electrophilicity in describing the biological activity has been enhanced to a reasonable height with the support of number of non-hydrogenic atoms (N_{NH}) as is delineated in the present investigation.

5. Conclusion

The number of atoms in a molecule may be a valid descriptor in the quantitative structure–activity relationship (QSAR) parlance. This descriptor alone is reasonably effective in explaining the biological activities of the estrogen derivatives unlike their testosterone counterparts. The use of this descriptor along with the electrophilicity index (ω) can explain the biological activities of a series of male and female hormones, viz., testosterone and estrogen derivatives in a satisfactory way.

Table 5
Electrophilicity index (ω) of estrogen derivatives with their experimental and calculated biological activities in terms of relative binding affinity (RBA) [24]

No.	Compounds	N_{NH}	Electrophilicity (ω) (eV)	Exptl RBA	Calc RBA	
					N_{NH}	N_{NH}, ω
16-α Substituted estradiol derivatives						
1	H	20	1.3583	2.00	2.04	2.03
2	CH_2Br	22	1.3685	1.97	1.76	1.76
3	$CH_2CH{=}CHCH_2OC_6H_5$	31	1.4588	0.85	0.48	0.48
17-α Substituted estradiol derivatives						
4	$C{\equiv}CMe$	23	1.3347	1.51	1.62	1.69
5	$CH_2CH{=}CH_2$	23	1.3586	1.26	1.62	1.65
6	C_6H_5	26	1.4573	1.08	1.19	1.12
7	$CH_2C_6H_5$	27	1.5526	0.63	1.05	0.85
8		26	1.4932	0.92	1.19	1.07
11-β,16-α,17-α Substituted estradiol derivatives						
	11-β	16-α	17-α			
9	Et	H	$C{\equiv}CH$	24	1.3712	1.94	1.47	1.51
10	Et	OH	Me	24	1.3898	1.93	1.47	1.48
11	H	OH	C_6H_5	27	1.3703	0.90	1.05	1.12
12	OMe	OH	C_6H_5	29	1.3317	0.70	0.76	0.93

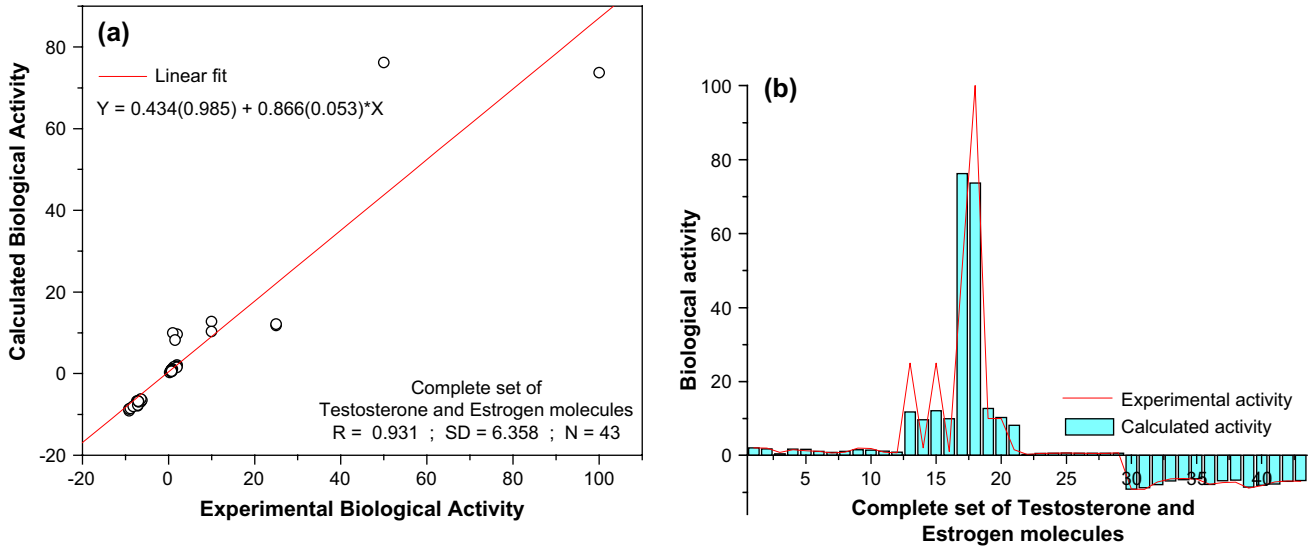


Fig. 1. (a) Calculated versus experimental biological activities and the (b) histogram representation of the same of all the testosterone and estrogen derivatives taken together.

Acknowledgements

We thank BRNS, Mumbai for financial assistance.

References

[1] J. Henderson, J. Endocrinol. 184 (2005) 5–10.
[2] Wikipedia–The Free Encyclopedia http://en.wikipedia.org/wiki/Main_Page.
[3] C.C. Chatterjee, Human Physiology, vol. II, Medical Allied Agency, Kolkata, 2003, 4/195–4/289.
[4] C. Hansch, P.P. Maloney, T. Fujita, R.M. Muir, Nature 194 (1962) 178.
[5] R. Franke, Theoretical Drug Design Methods, Elsevier, Amsterdam, NY, 1984, p. 115.
[6] M. Karelson, V.S. Lobanov, A.R. Katritzky, Chem. Rev. 96 (1996) 1027.
[7] M.T.D. Cronin, Curr. Opin. Drug Discov. Devel. 3 (2000) 292.
[8] C. Hansch, A. Kurup, R. Garg, H. Gao, Chem. Rev. 101 (2001) 619.
[9] R.G. Parr, W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, Oxford, 1989.
[10] P. Geerlings, F. De Proft, W. Langenaeker, Chem. Rev. 103 (2003) 1793.
[11] P.K. Chattaraj, S. Nath, B. Maiti, Reactivity descriptors, in: J. Tollenaere, P. Bultinck, H.D. Winter, W. Langenaeker (Eds.), Computational Medicinal Chemistry for Drug Discovery, Marcel Dekker, New York, 2004, pp. 295–322, Chapter 11.
[12] R.G. Parr, L.v. Szentpaly, S. Liu, J. Am. Chem. Soc. 121 (1999) 1922.
[13] P.K. Chattaraj, U. Sarkar, D.R. Roy, Chem. Rev. 106 (2006) 2065.
[14] P.K. Chattaraj, B. Maiti, U. Sarkar, J. Phys. Chem. A 107 (2003) 4973.
[15] R. Parthasarathi, J. Padmanabhan, V. Subramanian, U. Sarkar, B. Maiti, P.K. Chattaraj, Internet Electron. J. Mol. Des. 2 (2003) 798.
[16] U. Sarkar, D.R. Roy, P.K. Chattaraj, R. Parthasarathi, J. Padmanabhan, V. Subramanian, J. Chem. Sci. 117 (2005) 599;
D.R. Roy, U. Sarkar, P.K. Chattaraj, A. Mitra, J. Padmanabhan, R. Parthasarathi, V. Subramanian, S. Vandamme, P. Bultinck, Mol. Div. 10 (2006) 119.
[17] J. Padmanabhan, R. Parthasarathi, V. Subramanian, P.K. Chattaraj, Chem. Res. Toxicol. 19 (2006) 356.
[18] D.R. Roy, R. Parthasarathi, B. Maiti, V. Subramanian, P.K. Chattaraj, Bioorg. Med. Chem. 13 (2005) 3405.
[19] D.R. Roy, R. Parthasarathi, V. Subramanian, P.K. Chattaraj, QSAR Comb. Sci. 25 (2006) 114.
[20] R. Parthasarthi, V. Subramanian, D.R. Roy, P.K. Chattaraj, Bioorg. Med. Chem. 12 (2004) 5533.
[21] R.E. Counsell, P.D. Klimstra, F.B. Colton, Endocrinology 27 (1962) 248.
[22] G. Sala, G. Baldratti, Proc. Soc. Exp. Biol. Med. (NY) 95 (1957) 22.
[23] R.D. Cramer III, D.E. Patterson, J.D. Bunse, J. Am. Chem. Soc. 110 (1988) 5959.
[24] H. Gao, J.A. Katzenellenbogen, R. Garg, C. Hansch, Chem. Rev. 99 (1999) 723.
[25] V. Szymanski, W.R. Müller, J.V. Knop, N. Trinajstić, Int. J. Quantum Chem. S20 (1984) 173.
[26] M. Olaha, C. Bologaa, T.I. Oprea, J. Comp. Aid. Mol. Des. 18 (2004) 437.
[27] M.J. Frisch, Gaussian 03, Revision B.03, Gaussian, Inc, Pittsburgh, PA, 2003.
[28] P.P. Singh, H.K. Srivastava, F.A. Pasha, Bioorg. Med. Chem. 12 (2004) 171; R.G. Parr, P.K. Chattaraj, J. Am. Chem. Soc. 113 (1991) 1854.