  Bull. Chem. Soc. Jpn., 60, 1125—1128 (1987)

A Molecular Orbital Study of the Electronic Structure of Some 6-Substituted Uracils

Paul George ABDUL-AHAD and Graham Alan WEBB†,*
Scientific Research Council, Petroleum Research Center, P.O. Box 10039, Baghdad, Iraq
†Department of Chemistry, University of Surrey, Guildford, Surrey, U.K.
(Received March 19, 1986)

Semiempirical molecular orbital calculations are performed on a series of 28 6-[2-(substituted phenyl)-hydrazino]uracils. Approximate ground state electronic properties are obtained for each molecule studied as well as the electrostatic potentials. First and second order interaction energies are calculated for some selected point charges in the vicinity of the parent molecule. The best correlations obtained between the biological activity and the electronic and steric indicies are considered in a mechanistic context.

A number of 6-substituted uracils are reported to be potent inhibitors of *E. coli* thymidine phosphorylase.[1] Some of the 6-[2-(substituted phenyl)hydrazino]uracils are found to inhibit the replication of the specific enzyme DNA polymerase III of *Bacillus subtilis* by forming strong reversible complexes with the template primer DNA and the enzyme. *Bacillus subtilis* has three DNA polymerases. Of these polymerase I is involved in DNA repair, while polymerase III is necessary for DNA replication.[2]

The activity of the 6-(2-Phenylhydrazino)uracils indicates that most of the enzyme-inhibitor binding energy results from an interaction of the enzyme with the phenyl ring.[3] The phenyl ring-enzyme binding may involve van der Waals or possibly charge-transfer interactions with the amino acid. The mechanism of complex formation appears to involve hydrogen bonding of the uracil and its NH moiety with cytosine in the DNA template and binding of the phenyl group with a hydrophobic site on the enzyme.[4]

Table 1. Calculated Indices and Observed Activity of 6-(2-Phenylhydrazino)uracils

No.	Substituent R	pK_{50}[a]	B.E. / eV	m.v. / Å^3	$E_{C_{17}}^{BIND}$ / eV
1	H	5.8538	−373.35	186.86	−95.6866
2	4-OH	6.3979	−383.24	194.82	−95.7966
3	4-Me	6.3010	−407.37	203.37	−95.6514
4	4-F	5.5228	−375.38	191.42	−95.8183
5	4-Cl	6.0457	−372.28	200.39	−95.6941
6	4-NH_2	5.1487	−393.30	198.09	−95.7540
7	4-NO_2	4.5800	−406.62	209.38	−95.7754
8	3-Me	5.6020	−407.38	203.31	−95.2498
9	3-F	6.1549	−375.43	191.43	−95.2096
10	3-Cl	5.9586	−372.25	200.41	−95.1003
11	2-Me	3.6307	−407.41	202.86	−103.352
12	2-Et	3.2502	−423.97	214.09	−102.988
13	3,4-Me_2	5.8538	−441.22	218.70	−95.2252
14	3,4-Cl_2	6.2218	−371.64	213.72	−95.1259
15	3-Cl, 4-Me	6.0969	−406.39	216.36	−95.0642
16	3-Me, 4-OH	4.2907	−417.29	211.19	−95.3437
17	3-Cl, 4-OH	5.0555	−382.45	208.30	−95.1985
18	4-Br	6.1549	−375.61	204.55	−95.6498
19	3-Br	5.8538	−373.59	204.56	−95.4614
20	2-F	5.1426	−373.41	191.27	−102.899
21	4-Et	5.7958	−407.38	203.43	−95.2410
22	2-Me, 4-OH	6.3010	−417.30	211.21	−95.3565
23	3-Me	5.7958	−440.82	218.44	−95.6603
24	3,5-Me_2, 4-OH	3.9065	−451.33	226.94	−95.2773
25	4-CO_2H	4.0087	−416.20	213.44	−95.6931
26	4-SO_2Me	3.6798	−424.19	236.20	−95.6745
27	4-NHAc	3.2518	−459.82	231.66	−95.6889
28	4-*n*-Bu	2.6382	−508.91	252.91	−95.6588

a) pK_{50} is the negative logarithm of the molar concentration required to achieve 50% inhibition of enzyme activity.

The rotation of the phenyl N and hydrazino N–N bons are the basis for the formation of an active inhibitory conformation of the drug molecules. With a view to clarifying some of these assumptions the present work employs semiempirical molecular orbital (MO) methods in an attempt to interpret the mechanism of the action of 6-substituted uracils.

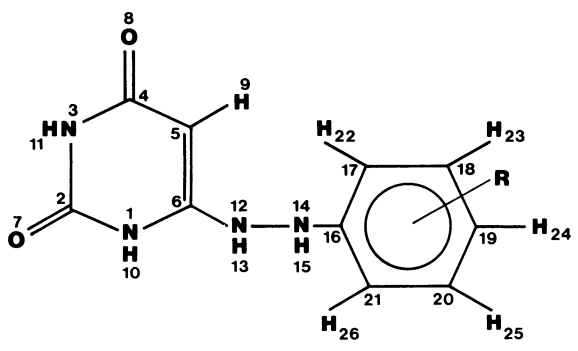
Method

The atom numbering scheme used in the present study is given in Structure I. The series of 28 6-[2-(substituted phenyl)hydrazino]uracils presently considered are given in Table 1 together with their calculated MO indices and observed activity data.[4] They are potent inhibitors of *Bacillus subtilis* DNA polymerase III. The inhibitory activities have been reported elsewhere[4] and are expressed as pK_{50}, which is the negative logarithm of the molar concentration required to achieve 50% inhibition of enzyme activity. Inactive compounds are excluded from the present study.

Two possible rotations about the phenyl N and N–N bonds are expected in the formation of an active inhibitory conformation of the drug molecule. The starting point geometry is obtained from reference.[5] This geometry is optimized using the program "GEOMIN" (QCPE 312), with the CNDO/2 option, to obtain a more reliable estimate of conformation in solution. Standard bond lengths and bond angles[6] are used for the substituent groups in the phenyl ring. The optimized bond lengths and angles are reported in Table 2.

The electrostatic potential (e.p.) of the drug molecule is calculated from the CNDO/2 wave functions using method III of reference,[7] in which monocentric differential overlap is retained. To assist in determining the most favorable binding position in the vicinity of the substrate an in-plane electrostatic contour map is plotted for the 6-(2-Phenylhydrazine)-uracils and given in Fig. 1. The lower the e.p. at a given position, the more energetically favorable is the placement of an electropositive reagent group at that position. The map is obtained using the simple plot program package of the University of Manchester.

First the second order interaction energies with each of five hypothetical point charges, placed at different positions in the molecular plane, are calculated. Positions P_1 and P_2 are chosen to coincide with the e.p. minima at -130 and -124 kcal mol^{-1} in the vicinity of O_7 and O_8, respectively, of the pyrimidine ring as shown in Fig. 1. P_3, P_4, and P_5 are positions which would be occupied by the O_2, N_3, and N_4–H atoms of a receptor cytosine group, assuming that these atoms become linked by hydrogen bonds to



Structure 1

Table 2. Geomin Optimized Bond Lengths, Bond Angles, and Dihedral Angles

Bond lengths/l/Å			
N_1–C_2	1.388	C_2–O_2 =C_4–O_8	1.277
C_2–N_3	1.379	N_1–H_{10} =N_3–H_{11}	1.067
N_3–C_4	1.391	N_{12}–H_{13} =N_{14}–H_{15}	1.077
C_4–C_5	1.435	C_{18}–C_{19} =C_{19}–C_{20}	1.384
C_5–C_6	1.341	C_{20}–C_{21}	1.384
C_5–H_9	1.112	C_{16}–C_{17}	1.397
C_6–N_{12}	1.402	C_{17}–C_{18}	1.382
N_1–N_{14}	1.342		

Bond angles/ϕ/°			
N_3–C_2–N_1	115.6	H_{13}–N_{12}–C_6	108.6
C_4–N_3–C_2	125.6	N_{14}–N_{12}–C_6	112.4
C_5–C_4–N_3	115.4	H_{15}–N_{14}–N_{12}	105.8
C_6–C_5–C_4	119.1	C_{16}–N_{14}–N_{12}	116.1
O_7–C_2–N_3	122.8	C_{17}–C_{16}–N_{14}	119.9
O_8–C_4–C_5	126.5	C_{18}–C_{17}–C_{16}	121.3
H_9–C_5–C_4	118.9	C_{19}–C_{18}–C_{17}	120.3
H_{10}–N_1–C_2	115.3	C_{20}–C_{19}–C_{18}	119.0
H_{11}–N_3–C_4	117.4	C_{21}–C_{20}–C_{19}	120.6
N_{12}–C_6–N_1			

Dihedral angles/ϕ/°	
C_4–N_3–C_2–N_1	346.4
C_5–C_4–N_3–C_2	5.35
C_6–C_5–C_4–N_3	359.2
O_7–C_2–N_3–C_4	169.7
O_8–C_4–C_5–C_6	178.7
H_9–C_5–C_4–N_3	180.7
H_{10}–N_1–C_2–N_3	163.7
H_{11}–N_3–C_4–C_5	171.9
N_{12}–C_6–N_1–C_2	164.0
H_{13}–N_{12}–C_6–N_1	44.9
N_{14}–N_{12}–C_6–N_1	168.2
H_{15}–N_{14}–N_{12}–C_6	131.5
C_{16}–N_{14}–N_{12}–C_6	250.6
C_{17}–C_{16}–N_{14}–N_{12}	210.0
C_{18}–C_{17}–C_{16}–N_{14}	183.7
C_{14}–C_{18}–C_{17}–C_{16}	359.6
C_{20}–C_{19}–C_{18}–C_{17}	359.9
C_{21}–C_{20}–C_{19}–C_{18}	0.27

N_{12}–H_{13}, N_1–H_{10}, and O_7, respectively, of the substrate. These hypothetical hydrogen bonds between substrate and pyrimidine template are shown in Fig. 1 of Ref. 4.

The first order interaction energy, with a point charge of unit magnitude, is simply the electrostatic potential. The second order interaction energy is calculated using uncoupled Hartree-Fock perturbation theory, with retention of monocentric differential overlap.[8]

Experimental

The calculations are performed on the CDC 7600 computer of the University of Manchester, using the "CNDO/2-3R" MO program (QCPE 261). The conformational analyses were performed on an ICL 2903 computer of the Iraqi Science Research Council.

The CNDO/2 wavefunctions are used to calculate all electronic reactivity indices. As well as the first and second order interaction energies, various simple indices are calculated. These may be calculated with a single atomic center, a pair of centers, or the whole molecule.

The first category is represented by the energy of binding of atom A to the rest of the molecule, $E_A^{BIND}=\sum_{B\neq A} E_{A-B}$, where E_{A-B} is the bond energy between centers A and B, calculated using the energy breakdown procedure.[9] E_{A-B} itself is used as an index of bond strength, this being a representative of the second category mentioned above.

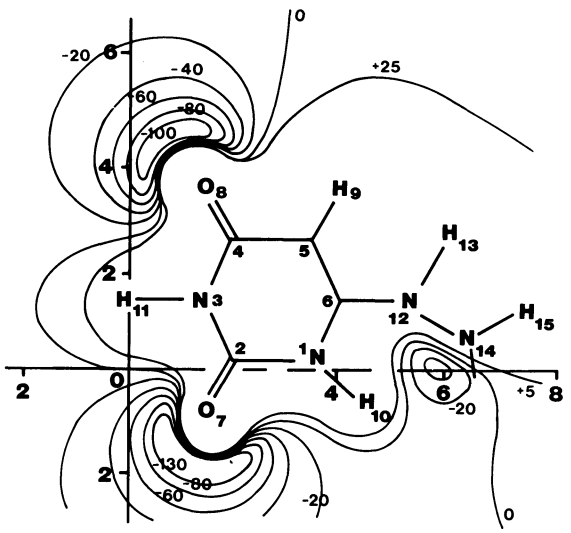


Fig. 1. The CNDO/2 e.p. contour map for 6-(2-phenylhydrazino)uracils (all values are in kcal mol^{-1}).

The following properties of the molecule as a whole are used as indices. The binding energy, B.E. defined as the difference between the CNDO/2 total energy of the molecule and the summed valance electronic energies of the isolated component atoms; the molecular volume, m.v., calculated by numerical integration of the solid produced by centering a sphere of the appropriate van der Waals radius[10] on each atomic center in the molecule. In calculating m.v., a rectangular solid is defined such that all of its faces touch the molecular profile. The solid is segmented into cubes of side 0.1 Å. It is then determined which of these cubes lie within the molecular profile, and a simple counting procedure is used to obtain the molecular volume.

Results and Discussion

The electronic and molecular indices relevant to the best regressions are given in Table 1. The most significant relationships, given in Table 3, have been obtained by searching the regression of inhibitive potency on each single and all pair contributions for eight parameters. The likelihood of random correlation increases with the number of regressions searched, therefore, caution is exercised in the interpretation of the statistical parameters.[11]

The positive coefficient associated with B.E. in regression (1), indicates that the less stable the molecule the higher the drug activity. The negative coefficient of the m.v. indicates that as the molecular size increase, it becomes less active. Since B.E. and m.v. have a substantial negative intercorrelation, it is difficult to decide which, if either, of these is a causal quantity.

The two-variable regressions (3) and (4) each contain an electronic index, as well as a size index. The appearance of the electronic properties of C_{17} tends to confirm the importance of the electronic structure of the phenyl ring. It is possible that there may be a receptor fragment located in this vicinity, in

Table 3. The Best Simple and Two-Variable Regressions Obtained for 6-(2-Phenylhydrazino)uracils

No.	Best simple regressions				No.	Best two variable regressions			
	pK_{50}[a] =	r[b]	s[c]	p[d]/%		pK_{50}[a] =	r[b]	s[c]	p[d]/%
1	0.022(±0.004)B.E. +14.48(±2.01)	0.672	0.85	99.9	3	0.233(±0.05)E_{C17}^{BIND} −0.0541(±0.009)m.v. +39.1(±5.9)	0.807	0.69	>99.9
2	−0.048(±0.01)m.v. +15.41(±2.4)	0.641	0.88	99.9	4	0.190(±0.05)E_{C17}^{BIND} +0.023(±0.004)B.E. +32.95(±5.8)	0.785	0.72	>99.9

a) Negative logarithm of the molar concentration required to achieve 50% inhibition of enzyme activity. b) Correlation coefficient. c) Standard error of estimate. d) Statistical significance determined from an F test.

the drug receptor complex. The positive coefficient associated with $E_{C_{17}}^{BIND}$ which is the only one of the calculated binding energies to enter the regressions, indicates that the more reactive this position the higher the drug activity. It appears that none of the first and second order interaction energies, for the chosen points in the molecular environment, show statistically significant correlations with pK_{50}. Consequently doubt is cast on the possible role of these energies in determining possible receptor binding sites at the chosen positions. Perhaps the electronic structure of the phenyl ring is critical in determining the biological activity of the 6-(2-phenylhydrazino)-uracils studied.

We conclude that, in order to design a more potent compound, with the same basic structure, it is apparently necessary to reduce the substituted molecular size as much as possible, whilst maximizing the reactivity of C_{17}.

References

1) B. R. Baker and W. Rzeszolarski, *J. Med. Chem.*, **10**, 1109 (1967).
2) K. B. Gass and N. R. Cazzarelli, *J. Biol. Chem.*, **248**, 7688 (1973).
3) G. E. Wright and N. C. Brown, *J. Med. Chem.*, **17**, 1277 (1974).
4) G. E. Wright and N. C. Brown, *J. Med. Chem.*, **20**, 1181 (1977).
5) L. L. Coulter and N. R. Cazzarelli, *J. Mol. Biol.*, **91**, 329 (1975).
6) M. S. Gordon and J. A. Pople, *J. Chem. Phys.*, **49**, 4643 (1968).
7) C. Giessner-Prettre and A. Pullman, *Theor. Chim. Acta*, **37**, 335 (1975).
8) P. W. Langhoff, M. Karplus, and R. P. Hurst, *J. Chem. Phys.*, **44**, 565 (1966).
9) J. A. Pople and D. L. Beveridge, "Approximate Molecular Orbital Theory," McGraw-Hill, New York (1970).
10) A. Bondi, *J. Phys. Chem.*, **68**, 44 (1964).
11) J. Topliss, *J. Med. Chem.*, **22**, 1238 (1979).