Pergamon

Bioorganic & Medicinal Chemistry, Vol. 4, No. 10, pp 1673–1677, 1996

PII: S0968-0896(96)00189-7
0968-0896/96 $15.00+0.00

Prediction of Relative Potency of Ketone Protease Inhibitors using Molecular Orbital Theory

A. W. Edith Chan[a,b,*] and Julian M. C. Golec[a]
[a]*Hoechst Marion Roussel, Kingfisher Drive, Covingham, Swindon, SN3 5BZ, U.K.*
[b]*Selectide, 1580 E. Hanley Blvd, Tucson, AZ 85737-9525, U.S.A.*

Abstract—Molecular orbital calculations were carried out on a series of model ketonic protease inhibitors. A comparison of the LUMO energy of the ketones in a variety of model heterocyclic ketone protease inhibitors shows a correlation with the electrophilicity of the carbonyl, and the σ_I experimental data. It is also observed that the more negative charge on the nitrogen atom in the heterocyclic ring the greater its potential as a hydrogen bond acceptor. The results of this study provide a simple means of predicting relative inhibitor potency and is therefore of use both to medicinal chemists designing protease inhibitors and in QSAR studies. Copyright © 1996 Elsevier Science Ltd

Introduction

The development of selective inhibitors to cysteine and serine proteases is of continuing interest to the pharmaceutical industry. Presently, thrombin[1] and human leukocyte elastase[2] are amongst the most actively studied serine proteases, while interleukin-1β convertase, a cysteine protease,[3] is receiving escalating attention owing to its potential role in inflammatory diseases and apoptosis.

The mechanism by which serine and cysteine proteases cleave peptide bonds is broadly similar. Potent reversible inhibitors have been made by replacing the substrate's scissile bond by an aldehyde.[4–7] Replacement of this bond by an alkyl ketone, however, is not generally sufficient to give satisfactory inhibition. Inhibition is usually achieved by the activation of the carbonyl in some way. This has been done by adding electron withdrawing groups, such as α-trifluoromethyl, α-carbonyl, or an α-heterocyclic group. Of primary interest here is the relationship between different α-keto heterocyclic groups and enzyme inhibition.

Edwards et al. chose benzoxazole as the activating heterocyclic group for inhibition of porcine pancreatic elastase[5] on the basis of its electron withdrawing effect on the adjacent carbonyl group and its potential ability to stabilize the hemiketal adduct through formation of a hydrogen bond with the protonated active site His-57 (see Fig. 1). The resulting inhibitors were more potent than the equivalent aldehyde and similar in potency to the analogous trifluoromethyl ketone. The expected hydrogen bond between the benzoxazole nitrogen and the protonated active site His-57 was consistent within the X-ray structure of the enzyme–inhibitor complex.

More recently a diverse series of α-heterocycles were evaluated for their ability to inhibit human neutrophil elastase (HNE).[6a] An attempt was made to correlate potency with the σ_I values or the electron withdrawing potential of the heterocycle. Unfortunately, this correlation was hampered by steric considerations and the ability of the heterocycle to hydrogen bond with the protonated active site His-57. Nevertheless, this work showed a good relationship between the electron withdrawing potential of the heterocycle and the relative potency against the enzyme.

In the design of heterocyclic ketone protease inhibitors, there is clearly a need for a simple means of assessing the relative merits of different heterocyclic groups. Experimentally derived σ_I values[8] have provided a means to this end. However, this data is limited. The aim of this study has been to take a series of α-keto heterocyclic compounds of the form: CH_3—(CO)—R, where R represents different substituent heterocyclic groups, and to find some theoretical parameters that correlate with the observed biological activity in ketone type inhibitors.[5–7] We have

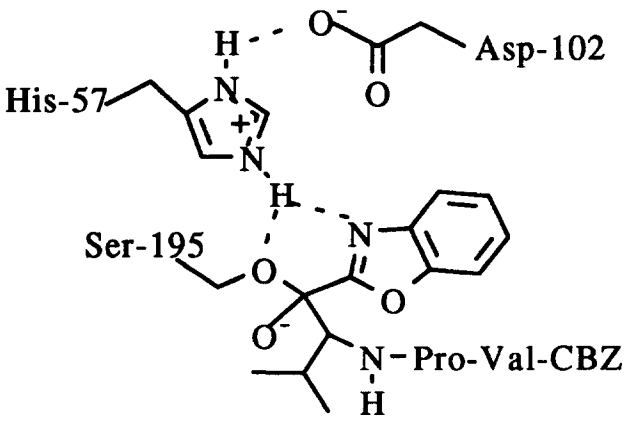


Figure 1. Covalent and hydrogen bonding interactions between the peptidyl α-ketoheterocycle Cbz—Val—Pro—Val—(CO)-benzoxazole and the catalytic site of a serine proteinase (taken from ref. 5).

employed molecular orbital theory at the Hartree–Fock (HF) level in order to predict the electrophilicity of the heterocyclic carbonyls (i.e. how activated the carbonyl is to nucleophilic attack)/the strength of the hydrogen bond of the heteroatom to the active site histidine. This work provides a simple means of predicting relative inhibitor potency as an alternative to using the limited number of existing experimentally derived σ_I values.[8]

Methods

Table 1 shows the 38 different R-groups of the CH_3—(CO)—R compounds considered in this study. The electronic calculations were done using MO theory at the HF level. The assumption made was that, in principle, nucleophilic attack[9,10] on the carbonyl can be thought of as a two-electron two-orbital bonding interaction between the π^* LUMO (lowest unoccupied molecular orbital) of the carbonyl and the lone pair HOMO (highest occupied molecular orbital) of the nucleophile (Fig. 2).[10] There are three main factors governing a nucleophilic addition reaction: the nature of the electrophile, nucleophile, and the solvent. In this study, the influence of the latter two factors was assumed to be constant. Perturbation molecular orbital theory[10] predicts that the lower the energy of the LUMO, the better the HOMO–LUMO interaction. It is therefore reasonable to compare the energy of the LUMO of all the compounds studied as a means of measuring the strength of the electrophilicity of the carbonyl. LUMO energies do not have the same status as those of HOMOs as they are virtual orbitals rather than orbitals occupied by spin, and may vary wildly with basis set. Here we restrict ourselves to one specific level of calculation so that at least in a comparative sense the energies are meaningful. Comparison of the carbonyl LUMO energies of the model compounds will show the relative electron withdrawing effect of the individual R-groups. In addition, we have used the electrostatic charge on the nitrogen atom of the heterocycle to quantify its strength as a hydrogen bond acceptor.[11] The electrostatic charge calculation was based on a set of atomic charge fits to the molecular electrostatic potential.[12]

The geometries of the model ketones presented in Table 1 were fully optimized using ab initio HF with the 3–21G* basis set.[12,13] The single point energy was calculated for each optimized compound using the 6–31G* basis set.[14]

The model inhibitors studied, CH_3—(CO)—R, could be divided into two categories according to the nature of the R group: noncyclic and heterocyclic systems. Table 1 lists the models and their respective calculated LUMO energy (eV). It is well-established from the first order perturbation molecular orbital theory that the frontier molecular orbitals (FMO) of a methyl group can simulate the FMO of an alkyl group or a simple peptide.[10] Moreover, the molecules studied are simple models of peptidyl inhibitors, allowing us to compare the important electronic properties of the R-groups using the same basis set in a reasonable amount of CPU time.

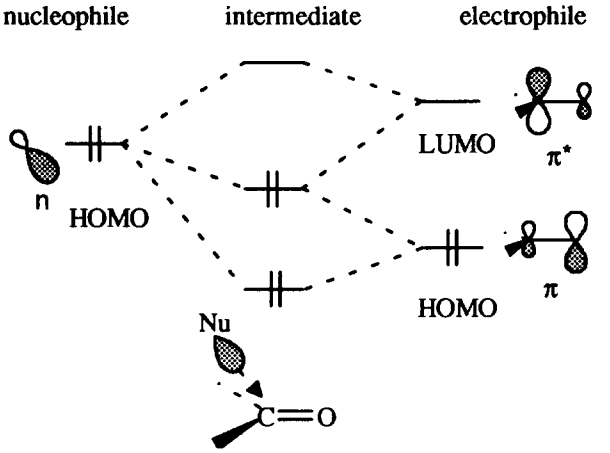


Figure 2. MO involved in a nucleophilic addition reaction. Note: the π HOMO of the carbonyl is not considered in this study since the energy of the LUMO (π^*) is the major factor governing the electophile by the first perturbation theory.[9,10]

The LUMO of all the models have π^* character localized at the C atom of the carbonyl as drawn in Figure 2. In the models with heterocyclic R groups, another important orbital for consideration is the heteroatom's lone pair. This lone pair orbital could potentially take part in hydrogen bond formation. We have assumed that the more negative the nitrogen atom, the better its hydrogen bond acceptor ability.

Results and Discussion

First we observed that the LUMO energies correlate well with the previously reported σ_I values.[8] For example, the values of σ_I for the non-heterocyclic carbonyls[5–8] R = Me **2**, H **1**, $CONH_2$ **8**, CO_2Et **7**, COMe **4** are $-0.01 < 0.00 < 0.28 < 0.30 = 0.30$, where the bigger the number, the more electron withdrawing that group is. The corresponding LUMO energies are $4.43 > 4.25 > 2.50 > 2.35 > 2.02$, respectively, where the lower the LUMO value, the better the interaction of the frontier orbitals. This correlation is also observed in the heterocyclic systems. For example, the values of σ_I for phenyl **31**, pyridine **35**, thiophene **14**, pyrimidine **33**, pyridazine **37**, thiazole **18**, benzoxazole **11**, and tetrazole **26** are $0.12 < 0.18 < 0.19 < 0.23 < 0.26 < 0.34 < 0.41 < 0.49$, respectively, while the corresponding LUMO energies are $2.48 > 2.20 > 2.18 > 1.99 > 1.96 > 1.81 > 1.76 > 1.74$, respectively. This demonstrates that the LUMO analysis can give reliable electronic information where no experimental σ_I data is available.

The LUMO energies for molecules where R is a noncyclic group corresponded with the K_i values (inhibition constant—the smaller it is the better the inhibition of the enzyme by the molecule) obtained by Edwards et al.[6a] on the inhibition of HNE with the peptidyl inhibitors [Ph—CH_2—O—(CO)]—Val—Pro-

—Val—R. The K_i values for R=COOMe, $CONH_2$, CF_3, H, and CH_3 were 0.6, 1.8, 1.6, 41, and 8000 nM, respectively, whereas the LUMO energies for the corresponding model compounds were 2.32<2.50<3.09<4.25<4.43. The activity trend is consistent with the lowering of the LUMO energy, and therefore shows that the lower the energy of the carbonyl LUMO the more potent the carbonyl inhibitor.

The LUMO energies of the non-heterocyclic model ketones, R=COMe **4**, COEt **6**, COOMe **5**, COOEt **7**, $CONH_2$ **8**, CF_3 **3**, H **1**, CH_3 **2** are 2.02<2.07<2.32<2.35<2.50<3.09<4.25<4.43, respectively. The diketo compounds **4** and **6** have the lowest LUMO in this group. The π^* character in the LUMO of **4** and **6** is localized on the C atoms of both carbonyl groups and thus providing two possible centers for nucleophilic attack. As the alkyl chain gets longer (from methyl to ethyl), however, the LUMO energy becomes higher.

In the series where R is a heterocylic group, the LUMO energies for thiazole **18**, oxazole **17**, and imidazole **20** are 1.81<2.22 <2.68, respectively. The corresponding electrostatic charges on the nitrogen are −0.43, −0.53, and −0.53, respectively, where the hydrogen bond acceptor ability of nitrogen is the least in thiazole **18**. This trend is also observed in the benzothiazole **10**, benzoxazole **11**, benzimidazole **12**.

An important point to note is that the effect of all these inhibitors is likely to be different for different proteases, being governed by the structure of the active site of the proteases. For example, Tsutsumi[7] observed the K_i values for peptidyl benzothiazole, benzoxazole, thiazole, and imidazole in prolyl endopeptidase (PEP) to be 4.0, 5.6, 5.0, and 9.0 nM, respectively, while Edwards[6a] observed the K_i values for peptidyl benzothiazole, benzoxazole, thiazole, and oxazole in to be 25.0, 3.0, 270.0, and 28.0 nM, respectively. The LUMO energies (and the electrostatic charge on nitrogen) for

Table 1. LUMO energies and electrostatic charges of CH_3—CO—R

	R group	LUMO (eV)	Charges β-N	σ_I5,6		R group	LUMO (eV)	Charges β-N	σ_I
1	-H	4.25		0.00	**2**	$-CH_3$	4.43		−0.01
3	$-CF_3$	3.09		—	**4**	-COMe	2.02		0.30
5	-COOMe	2.32		—	**6**	-COEt	2.07		—
7	-COOEt	2.35		0.30	**8**	$-CONH_2$	2.50		0.28
9		1.41	−0.66	—	**10**		1.63	−0.49	0.37
11		1.76	−0.62	0.41	**12**		2.04	−0.63	0.32
13		2.58		0.17	**14**		2.18		0.19
15		2.62	−0.66	—	**16**		2.36	−0.47	—
17		2.22	−0.53	0.38	**18**		1.81	−0.43	0.34
19		2.62	−0.73	—	**20**		2.68	−0.53	0.27
21		1.92	−0.33	—	**22**		3.02		0.17
23		1.49	−0.72	0.49	**24**		2.10	−0.66	—
25		1.15	−0.47	—	**26**		1.74	−0.40	0.49
27		1.90	−0.38	—	**28**		1.82	−0.49	—
29		2.54	−0.64	—	**30**		1.91	−0.40	—
31		2.48		0.12	**32**		1.40	−0.94	—
33		1.99	−0.82	0.23	**34**		1.92		—
35		2.20	−0.65	0.18	**36**		2.20		—
37		1.96	−0.26	0.26	**38**		1.59	−0.44	0.37

benzothiazole **10**, benzoxazole **11**, thiazole **18**, oxazole **17**, and imidazole **20** are 1.63(−0.49), 1.76(−0.62), 1.81(−0.43), 2.22(−0.53), and 2.68(−0.53), respectively. This suggests that in the first case, PEP, the activation of the carbonyl weighs heavier than the hydrogen bond factor, while in the second case, HNE, the hydrogen bond acceptor ability of the nitrogen atom plays a more important role. Nevertheless, both factors should contribute to the K_i values. Edwards[6b] has further synthesized a set of peptidyl benzothiazoles with both electron-donating and withdrawing substitutions at the 5 position of the benzothiazole for the inhibition of HNE. Their reported K_is did not correspond well with our calculated LUMO values. Primary examination of the X-ray crystal structure of the active site of HNE showed that there could be steric effect between the binding pocket and the substitution at the 5 position.

The LUMO energies of oxazoline **15** and thiazoline **16** are 2.62 and 2.36 eV, respectively. These are higher than the corresponding oxazole and thiazole, implying poorly activated carbonyls. However, in two different studies, peptidyl α-keto oxazoline[5] and thiazoline[6a] were found to be better protease inhibitors than the corresponding α-keto oxazole and thiazole. Edwards[5] proposed that this observation was due to the hydrogen bond with the active site histidine. Comparison of the volume and the surface area of the nitrogen lone pair orbital of oxazole (15.88 Å^3 and 65.12 Å^2, respectively) and oxazoline (19.37 Å^3 and 77.62 Å^2, respectively) shows that the nitrogen lone pair of the latter is much larger than that of the former. The increased orbital overlap provides better hydrogen bonding potential. The electrostatic charge on the nitrogen of oxazoline **15**, oxazole **17**, thiazoline **18**, and thiazole **16** are −0.66, −0.53, −0.47 and −0.43, respectively, confirming the view that a reduced heterocyclic system can give a better hydrogen bond.

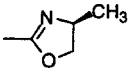
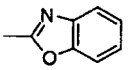
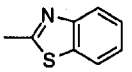
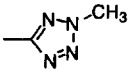
Comparison of furan **13**, oxazole **17**, 1,3,4-oxadiazole **21**, and 1,2,4-oxadiazole **23**, and oxatriazole **25** demonstrate the effect of the additional nitrogen atoms into an oxygen containing 5-membered ring system. The LUMO energies for the above series are 2.58 > 2.22 > 1.92 > 1.49 > 1.15, respectively. Thus addition of more heteroatoms to the ring system lowers the LUMO energy and further activates the carbonyl to nucleophilic attack. This provides a way of increasing the electron withdrawing power of the conjugate π system. This trend is also seen in the thio-series (Table 1).

The α-keto imidazole **20**, triazole **24**, and tetrazole **26** in which the ring systems contain only nitrogens, have a similar trend to the series mentioned above. In general though, the LUMO energies of this series are higher than the other corresponding series, implying a poor carbonyl activating center. This suggests that in order to lower the energy of the LUMO and maintain the nitrogen's hydrogen bond acceptor ability, it is better to place two different heteroatoms in the ring.

The trends observed in six-membered ring systems are similar to those of the five-membered ones. The LUMO energy of α-keto phenyl **31**, 2-pyridine **35**, 3-pyridine **36**, 2- pyrimidine **33**, 5-pyrimidine **34**, 1,2,4-triazine **38**, and 1,3,5-triazine **32** are 2.48 > 2.20 = 2.20 > 1.99 > 1.92 > 1.59 > 1.40, respectively. The electrostatic charges on the nitrogen are in general more negative than those in the five-membered ring cases. The rings with a β-nitrogen have charges of −0.65, −0.62, and −0.93 for 2-pyridine **35**, 1,2,4-triazine **38**, and 1,3,5-triazine **32**, respectively. Unfortunately, there are few examples of protease inhibitors containing carbonyl groups activated by 6-membered rings. Tsutsumi[7] observed K_i values for the inhibition of PEP by peptidyl 2-pyridine, 3-pyridine, and 5-pyrimidine as 6.9, 2290.0, and 61.0 nM, respectively. The relative inactivity of the 3-pyridine may be explained by a higher LUMO value and its inability to hydrogen bond to the His at the active site.

We have analyzed (Table 2 and Fig. 3) the structure–activity relationship between ln (K_i), the calculated energy of the LUMO, and the calculated charge on the β-nitrogen in the cyclic systems[6a] by the partial least squares method (PLS).[15] With only two variables, LUMO energy and charge on the nitrogen, the statistical significance is 0.91, with cross validation (r(cv)^2) and fraction of variance (r^2) being 0.69 and 0.78, respectively. This shows that indeed we can be confident in using these parameters to rationalize the struc-

Table 2. Comparison of K_i, ln (K_i), σ_I, LUMO energy, and the electrostatic charge on N in the heterocyclic rings for peptidyl α-ketoheterocycles

	R	K_i[a] (ln (K_i))	σ_Ia	LUMO (eV)[b]	Charge on N[b]
15		0.55 (−0.60)	0.32	2.62	−0.66
19		1.2 (0.18)	—	2.62	−0.73
11		3.0 (1.10)	0.41	1.76	−0.62
10		25 (3.22)	0.37	1.63	−0.49
17		28 (3.33)	0.38	2.22	−0.53
29		49 (3.89)	—	2.54	−0.64
18		270 (5.60)	0.34	2.36	-0.43
30		410 (6.02)	—	1.91	−0.40

[a][Ph—CH_2—O—(CO)]—Val—Pro—Ala—R (tested compound, ref. 6).
[b]CH_3—CO—R (compound in calculations).

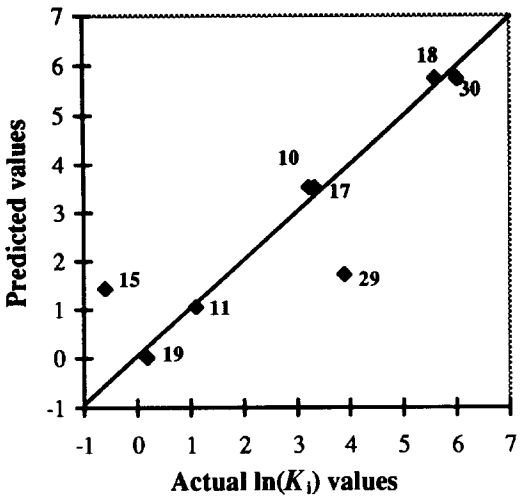


Figure 3. Actual versus predicted ln (K_i) values for the data set in Table 2 using PLS analysis.

ture–activity of these compounds and with other additional relevant parameters, a better QSAR model could be obtained for this kind of peptidyl compound.

Conclusion

Peptidyl α-keto heterocycles have found use as inhibitors of cysteine and serine proteases.[4–7] We have demonstrated that the calculated LUMO energy of the ketone in different α-keto model compounds and the calculated negative charge on the β-nitrogen of the activating heterocycles can be correlated to the experimental[5–7] inhibition constant of these molecules. On the basis of these correlations, we have tabulated the LUMO values of a large set of templates which can guide medicinal chemists in designing inhibitors for other proteases. Furthermore, these simple calculations may be used to give guidance to the likely activity of substituted heterocyclic groups. Of course, steric considerations[6b] between the potential inhibitor and the protease still need to be taken into account. However, where there is no structural information on the enzyme or experimental σ_I values for the heterocyclic rings in question, the LUMO values can still provide a useful starting point. This study shows how, with the right computer software and hardware a non-specialist can perform simple calculations.

Acknowledgement

Thanks go to Drs M. J. Batchelor, D. Bebbington, S. D. Jones, R. A. Laskowski, D. M. Proserpio, and W. G. Richards for helpful discussions.

References

1. (a) Wiley, R. A.; Rich, D. H. *Med. Chem Rev.* **1993**, *13*, 327. (b) Jakubowski, J. A.; Smith, G. F.; Sall, D. J. *Annu. Rep. Med. Chem.* **1992**, *27*, 99. (c) Stone, S. R. *Trends Cardiovas. Med.* **1995**, *5*, 134.

2. (a) Edwards, P. D.; Bernstein, P. R. *Med. Res. Rev.* **1994**, *14*, 127 and refs therein. (b) Stein, R. A.; Strimpler, A. M.; Edwards, P. D.; Lewis, J. J.; Manger, R. C.; Schwartz, J. A.; Stein, M. M.; Trainor D. A.; Wildonger, R. A.; Zottola, M. A. *Biochemistry* **1987**, *26*, 2682.

3. (a) Black, R. A.; Kronheim, S. R.; Sleath, P. R. *FEBS Lett.* **1989**, *247*, 386. (b) Kostura, M. J.; Tocci, M. J.; Limjuco, G.; Chin, J.; Cameron, P.; Hillman, A. G.; Chartrain, N. A.; Schmidt, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5227. (c) Dolle, R. E.; Hoyer, D.; Prasad, C. V. C.; Schmidt, S. J.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *J. Med. Chem.* **1994**, *37*, 563. (d) Kroemer, G.; Petit, P.; Zamzami, N.; Vayssiere, J.-L.; Mignotte, B. *FASEB J.* **1995**, *9*, 1277. (e) Kumar, S.; Harvey, N. L. *FEBS Lett.* **1995**, *375*, 169.

4. Mehdi, S. *Bioorg. Chem.* **1993**, *21*, 249 and refs therein.

5. Edwards, P. D.; Meyer, Jr E. F.; Vijayalakshmi, J.; Tuthill, P. A.; Andisik, D. A.; Gomes, B.; Strimpler, A. *J. Am. Chem. Soc.* **1992**, *114*, 1854.

6. (a) Edwards, P. D.; Wolanin, D. J.; Andisik, D. A.; Davis, M. W. *J. Med. Chem.* **1995**, *38*, 76. (b) Edwards, P. D.; Zottola, M. A.; Davis, M.; Williams, J.; Tuthill, P. A. *J. Med. Chem.* **1995**, *38*, 3972.

7. Tsutsumi, S.; Okonogi, T.; Shibahara, S.; Ohuchi, S.; Hatsushiba, E.; Patchett, A. A.; Christensen, B. G. *J. Med. Chem.* **1994**, *37*, 3492.

8. Taylor, P. J.; Wait, A. R. *J. Chem. Soc. Perkin. Trans. 2* **1986**, 1765.

9. (a) Bürgi, H.-B.; Dunitz, J. D.; Shefter, E. J. *J. Am. Chem. Soc.* **1973**, *95*, 5065; *Acta Crystallogr., Sect. B* **1974**, *30B*, 1517. (b) Bürgi, H.-B.; Lehn, J. M.; Wipff, G. *J. Am. Chem. Soc.* **1974**, *96*, 1956. (c) Bürgi, H.-B.; Dunitz, J. D.; Lehn, J. M.; Wipf, G. *Tetrahedron* **1974**, *30*, 1563. (d) Minato, T.; Fujimoto, H.; Fukui, K. *Bull. Chem. Soc. Jpn* **1978**, *51*, 1621. (e) Eisenstein, O.; Hoffmann, R. *J. Am. Chem. Soc.* **1981**, *103*, 4308.

10. Albright, T. A.; Burdett, J. K., Whangbo, M. H. In *Orbital Interactions in Chemistry*; Wiley-Interscience, 1985; Chapter 3, 10.5, and refs therein

11. Abraham, M. H.; Duce, P. P.; Prior, D. V.; Barratt, D. G.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1989**, 1355.

12. *Spartan 3.0* program running on a Indigo II workstation; distributed by Wavefunction, Inc., Irvine, CA 92715, U.S.A.

13. (a) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939. (b) Gorden, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797. (c) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; De Frees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039. (d) Dobbs, K. D.; Hehre, W. J. *J. Comput. Chem.* **1986**, *7*, 359.

14. (a) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, 56, 2257. (b) Hariharan, P. C.; Pople, J. A. *Chem. Phys. Lett.* **1972**, *16*, 217. (c) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gorden, M. S.; De Frees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654.

15. TSAR program from Oxford Molecular.

(*Received in U.S.A.* 29 *April* 1996; *accepted* 18 *June* 1996)