Pergamon

Bioorganic & Medicinal Chemistry, Vol. 3, No. 3, pp. 217–225, 1995
Copyright © 1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0968-0896/95 $9.50 + .00

0968-0896(95)00017-8

Quantitative Structure–Activity Relationships of Competitive Inhibitors of Phosphoenolpyruvate Carboxylase

Ricardo L. Mancera,[a] Ariadne G. Gómez[b] and Alejandro Pisanty*

Departamento de Física y Química Teórica, Facultad de Química, Universidad Nacional Autónoma de México (UNAM), Ciudad Universitaria, 04510 México, D.F., México

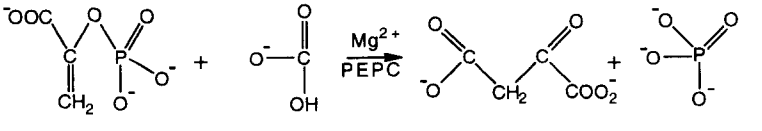
Abstract—The quantitative structure–activity relationships (QSAR) of all known competitive inhibitors of the enzyme phosphoenolpyruvate carboxylase from C_4 plants were investigated by means of molecular mechanics, the semiempirical quantum chemical methods MNDO and AM1, and the Hansch approach. In the case of phosphoenolpyruvate analogues, the hydrophobicity and steric impediment of the combined *cis* and *trans* substituents, the bond distance to the *cis* substituent along with its volume, dipole moment, the distance between the phosphorus and the carbonyl carbon, and the net electric charges on the phosphate and substituent groups are the main factors that govern their binding to the active site. For the phosphoglycolate analogues, the difference in the HOMO–LUMO energies, the magnitudes of their dipole moments and their non-polar surfaces, and the distance between the phosphorus and the carbonyl carbon are the variables that control their binding to the active site. These results, in conjunction with a discriminant analysis, also suggest that these inhibitors can actually be divided into two groups, according to the way they presumably interact with the active site.

Introduction

Phosphoenolpyruvate (PEP) carboxylase (PEPC) is an enzyme that plays an important role in the metabolism of C_4 and CAM plants by allowing them to fix CO_2 by catalyzing the β-carboxylation of PEP to produce oxaloacetate and inorganic phosphate,[1,2] as seen in Figure 1. This reaction constitutes the main step of the Hatch–Slack pathway. This metabolic pathway enables these plants to have a higher growing rate and biomass production. Among C_4 plants there are some species of economical importance like maize, sugar cane and sorghum. However, most of the weeds are also C_4 plants, so the development of specific growth inhibitors for these should be of great economical interest. PEPC has most often been the favorite target for this type of compound.

During the last few years, a large amount of information about the reaction mechanism of PEPC and the structure of its active site has been obtained, mainly by chemical modification studies and the use of PEP structural analogues as substrates and inhibitors. Chemical modification and kinetic studies of the enzyme have shown that cysteine, lysine, histidine and arginine residues are essential for the catalytic function.[2–6] Unfortunately, very little is known about the chemical environment and the precise role of these residues in the active site of the enzyme.

There exists a series of competitive inhibitors of PEPC that possess a close structural analogy to PEP, and that have been used to study the kinetic mechanism of this enzyme, the chemical nature of the amino acid residues involved in the binding of the substrate, and the topography and size of the active site. The names, chemical structures and inhibition constants (K_i) of these compounds[1,7–11] are shown in Figure 2.

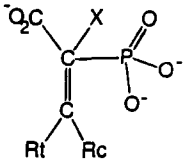


PEP Bicarbonate Oxaloacetate Phosphate

$\Delta G_r = -7.2$ *kcal/mol*

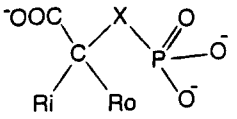
Figure 1. Reaction of PEP catalyzed by PEP carboxylase.

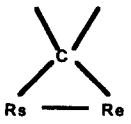
[a]Present address: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.
[b]Present address: Department of Chemical Engineering, University of Cambridge, Pembroke Street, Cambridge CB2 3RA, U.K.

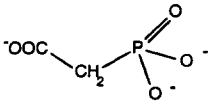


Rt = H X = O		
Rc	**Name**	**K_i (μM)**
Br	Z-Bromo-PEP	7
CH_3	Z-Methyl-PEP	17
Cl	Z-Chloro-PEP	64
F	Z-Fluoro-PEP	85
Rc = H X = O		
Rt	**Name**	**K_i (μM)**
CH_3	E-Methyl-PEP	110
CN	E-Cyano-PEP	1360
X = O		
Rt=Rc	**Name**	**K_i (μM)**
CH_3	Dimethyl-PEP	380
Rc = Rt		
X	**Name**	**K_i (μM)**
CH_2 (Rc=Rt=Cl)	3,3-dichloro-2-dihydroxiphosphinoil-methyl-2-propenoate (DCDP)	80
CH_2 (Rc=Rt=H)	PEP-Phosphonate	400
S (Rc=Rt=H)	Phosphoenoltiopyruvate (S-PEP)	2000

Figure 2(a).



X = O			
Ro	**Ri**	**Name**	**K_i (μM)**
CH_3	H	L-2-Phospholactate	100
H	H	Phosphoglycolate	200
CH_2-COO^-	H	Phosphomalate.	2700
 (Ro = Ri = CH2)		1-hydroxicyclopropanoic acid phosphate	7
X = CH_2			
Ro	**Ri**	**Name**	**K_i (μM)**
H	H	Phosphonopropionate	10000
X = S			
Ro	**Ri**	**Name**	**K_i (μM)**
H	H	Phosphotioglycolate	16000



Name	K_i (μM)
Phosphonacetate	2000

Figure 2(b).

Figure 2. PEP inhibitors studied in this paper. (a) Phosphoenolpyruvate analogs. (b) Phosphoglycolate analogs.

Several hypotheses have been postulated in connection with the different degrees of inhibition shown by the different above mentioned compounds, in view of their molecular characteristics. O'Leary[8] explains that the C–C–O bond angle has a very important effect on binding (compare the series phosphoglycolate, phospholactate, F-PEP, Br-PEP, and 1-HCP), since the latter is increased as the bond angle changes from about 109° (in 1-HCP) to about 120°. She also draws attention to the fact that any substitution of the oxygen of the phosphate ester[8,9] (with a $-CH_2-$ or a –S– group) results in a considerable decrease in the binding strength. This may be due to a hydrogen bonding between the oxygen and some group on the active site of the enzyme, or possibly to the variations in bond lengths, bond angles and/or electronic structure associated with these substitutions. Finally, she mentions[8] that it is not known whether large substituents on the double-bond end of PEP analogues can be accommodated. *Z*-Br-PEP, *Z*-Cl-PEP, and *Z*-F-PEP bind well to the enzyme, but phosphomalate (a phosphoglycolate analogue) cannot do so easily. González and Andreo[7] explain that substitutions with larger (e.g. –Br, –Cl or $-CH_2$) or more electronegative (e.g. –F, –Cl or –Br) groups result in a significant increase in the binding strength. They also mention[2] that, apparently, *trans* substitutions are responsible for the steric impediment that decreases the binding strength, while *cis* substitutions increase it. It is important to mention here that Izui *et al.*[12] found that near the active site of the PEPC from *E. coli* there exists a hydrophobic pocket large enough to allow the interaction with alkyl groups in a *cis* configuration, like ethyl, propyl and isopropyl groups, but not large enough to allow the interaction with an amyl group.

This paper reports a study of the quantitative structure–activity relationships of all the known competitive inhibitors of PEPC. Few of these compounds have been the subject of a thorough enzyme-kinetic study until now and there is not a large number of them, compared to the number of compounds usually available for QSAR studies; however, we have attempted to rationalize their observed biological activity, even at the price of only finding limited but still acceptable statistical quality in our results. To our knowledge, no systematic study of the relationship between the chemical structure and biological activity of PEPC modulators has ever been reported.

It should also be pointed out that PEPC has not yet been crystallized in such a way as to permit a high-resolution determination of its three-dimensional structure and/or of the active site structure. Therefore, no docking studies or similar refined modelling is possible at this stage. Our work does as much as is possible in the study of structure–activity relationships within the framework of the available kinetic and structural information, and the methods used. Our results can serve as encouragement and a first guide for further experimental work, and is actually already being used as such by experimental organic chemists and enzyme biochemists at UNAM.

Our goal was a first approximation to the problem of identifying the intrinsic molecular properties of the competitive inhibitors that are responsible for their different degrees of activity, as a result of the forces involved in their competitive (towards PEP) binding to the active site of PEPC. A first approach was followed using the Hansch method,[13-16] which makes use of hydrophobic (π), electronic (σ), steric (E_s) and dispersion (MR) parameters to quantitatively correlate activity to physicochemical descriptors. We then made use of molecular mechanics (MM) and the semiempirical quantum-chemical methods MNDO and AM1 to study the structural properties (interatomic distances, bond angles, dihedral angles, etc.) of the equilibrium geometries of the compounds, along with the electronic structure (net charges, orbital energies, HOMO and LUMO contributions, etc.) associated with these geometries.

Although the interaction with magnesium and, to a lesser extent, manganese ions has been identified as an important factor in PEPC action,[1,2,17,18] the mechanism of this interaction has not been established unambiguously in detail.[19–21] In this paper, the interaction with metal ions has not been considered because of the additional complexity and uncertainty it would introduce into the calculations, and in order to first study correlations between activity and the properties of the isolated molecules.

Results and Discussion

Hansch method

The method of Hansch was applied to describe the structural variations of only some of the PEP analogues, since substitutions of the oxygen of the phosphate ester cannot be accounted for with the available constants for substituents. For this reason, DCDP, PEP phosphonate and S-PEP were not included in this analysis. Phosphoglycolate analogues were not studied since they possess practically no structural variations on their substituents (either –H or $-CH_2$). The only significant correlation that was found is the following:

$$\log(1/K_i) = -1.510(\pm 0.397) + 1.132(\pm 0.436)E_s + 1.862 \quad (\pm 0.545)\pi \qquad (1)$$

$n = 7,\ r = 0.86297,\ r^2 = 0.74473,\ s = 0.47694,\ F = 5.83467\ (6.5165\%),\ r_c^2 = 0.5577$

where E_s is the steric Taft constant and π is the hydrophobic Hansch constant, in both cases for the combined *cis* and *trans* substituents. Unfortunately, the two variables involved are highly colinear, with $r_c^2 = 0.5577$, but the equation does show good correlation and is included for completeness. No correlation was found individually for any of the two parameters. The plot of equation 1 is shown in Figure 3.

Equation 1 is, unfortunately, unable to distinguish the effects of the *cis* and *trans* substituents; however, it reveals that the overall increases of hydrophobicity and steric impediment allow a much stronger interaction with the active site, possibly with a somewhat large and hydrophobic area.

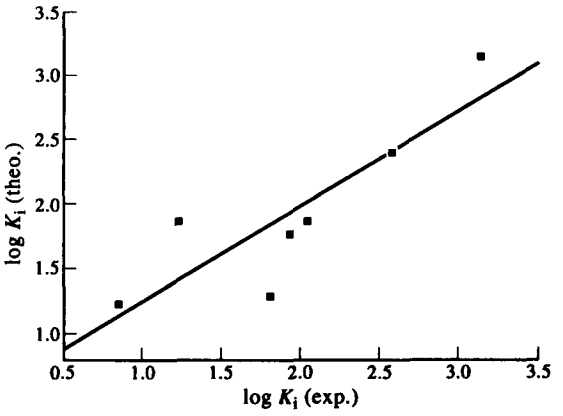


Figure 3. Test of fit of Hansch method for PEP analogs.

Molecular mechanics

The results of the MM calculations for the PEP analogues produced the following best correlations:

$$\log(1/K_i) = -4.380(\pm0.918) + 1.640\ (\pm0.623)\,C_3R_{cis} \quad (2)$$
$n = 9,\ r = 0.70535,\ r^2 = 0.49752,\ s = 0.53334,\ F = 6.93081\ (3.3783\%)$

$$\log(1/K_i) = -7.720(\pm3.719) + 1.378\ (\pm0.690)\,C_3R_{cis} + 0.028(\pm0.030)\,POC \quad (3)$$
$n = 9,\ r = 0.74869,\ r^2 = 0.56054,\ s = 0.53874,\ F = 3.82652\ (8.4872\%),\ r_c^2 = 0.1683$

$$\log(1/K_i) = -1.052(\pm2.644) + 2.003\ (\pm0.651)\,C_3R_{cis} - 0.021(\pm0.015)\,Tot \quad (4)$$
$n = 9,\ r = 0.78253,\ r^2 = 0.61235,\ s = 0.50599,\ F = 4.73887\ (5.8255\%),\ r_c^2 = 0.1745$

$$\log(1/K_i) = -3.918(\pm0.929) + 1.942\ (\pm0.627)\,C_3R_{cis} - 0.632(\pm0.461)\,C_3R_{trans} \quad (5)$$
$n = 9,\ r = 0.78570,\ r^2 = 0.61732,\ s = 0.50273,\ F = 4.83945\ (5.6041\%),\ r_c^2 = 0.1235$

$$\log(1/K_i) = -26.998(\pm14.203) + 1.900\ (\pm0.587)\,C_3R_{cis} + 0.189(\pm0.119)\,CCO \quad (6)$$
$n = 9,\ r = 0.80446,\ r^2 = 0.64716,\ s = 0.48273,\ F = 5.50241\ (4.3920\%),\ r_c^2 = 0.0771$

where C_3R_{cis} and C_3R_{trans} are the bond distances between C-3 and the first atom of the *cis* and *trans* substituents, respectively, *POC* is the bond angle P–O–C, *CCO* is the bond angle C–C–O$^-$, and *Tot* is the total van der Waals surface.

Correlations 2 to 6 reveal a predominance of the variable C_3R_{cis}, which accounts for 50% of the variance of the data. This suggests that the most important factor that contributes to a stronger binding to the active site is the increase in the distance between the *cis* substituent and the double-bond end of PEP analogues, as calculated by MM.

The results of the MM calculations for the phosphoglycolate analogues produced the following two best correlations:

$$\log(1/K_i) = -6.624(\pm1.334) + 0.436(\pm0.143)\,DM \quad (7)$$
$n = 6,\ r = -0.83661,\ r^2 = 0.69991,\ s = 0.70679,\ F = 9.32942\ (3.7865\%)$

$$\log(1/K_i) = -7.150(\pm0.495) + 0.826(\pm0.091)\,DM - 0.076(\pm0.015)\,SNP \quad (8)$$
$n = 6,\ r = 0.98502,\ r^2 = 0.97027,\ s = 0.25688,\ F = 48.95564\ (0.5126\%),\ r_c^2 = 0.6732$

where, in addition to the variables already defined, *DM* is the magnitude of the dipole moment and *SNP* is the saturated non-polar van der Waals surface. The corresponding plot of equation 8 is shown in Figure 4.

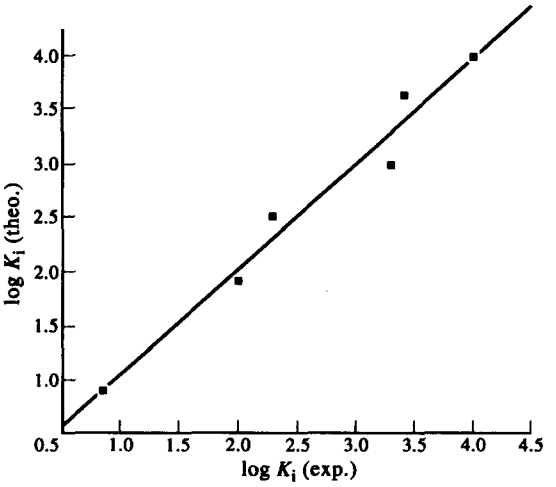


Figure 4. Test of fit for best fit of molecular mechanics results for phosphoglycolate analogs.

Equation 8 suggests that non-specific electrostatic interactions (as those being modelled by *DM*) are the most important forces involved in the binding of phosphoglycolate analogues to the binding site. On the other hand, since inhibition is increased with a decrease in the saturated non-polar surface of these compounds, it is possible to think that there are also hydrophobic-type forces involved in the binding to a non-polar area in the enzyme.

MNDO

The results of the MNDO calculations for the PEP analogues produced the following best correlations:

$$\log(1/K_i) = -18.250(\pm6.885) - 8.479(\pm3.778)\,qPO_4^{2-} + 0.142(\pm0.040)\,DM \quad (9)$$
$n = 10,\ r = 0.81362,\ r^2 = 0.66198,\ s = 0.51383,\ F = 6.85433\ (2.2455\%),\ r_c^2 = 0.1333$

$$\log(1/K_i) = -8.613(\pm2.331) + 0.601(\pm0.240)\,LH + 0.147(\pm0.038)\,DM \quad (10)$$
$n = 10,\ r = 0.83279,\ r^2 = 0.69354,\ s = 0.48925,\ F = 7.92086\ (1.5933\%),\ r_c^2 = 0.1575$

$$\log(1/K_i) = -15.668(\pm6.166) - 7.082(\pm3.399)\,qPO_4^{2-} + 0.065(\pm0.016)\,VR_{cis} \quad (11)$$
$n = 10,\ r = 0.84023,\ r^2 = 0.70598,\ s = 0.47922,\ F = 8.40405\ (1.3782\ \%),\ r_c^2 = 0.0685$

where, additionally, qPO_4^{2-} is the charge on the phosphate, *LH* is the difference of energy LUMO–HOMO, and VR_{cis} is the van der Waals volume of the *cis* substituent.

The best correlation (equation 11) shows that the decrease of the charge on the phosphate (representative of specific electrostatic interactions) and the increase of the van der Waals volume of the *cis* substituent produce a stronger binding to the active site. The other correlations seem to support the fact that specific and non-specific electrostatic interactions are the main factors governing active-site binding of these compounds.

The results of the MNDO calculations for the phosphoglycolate analogues produced the following best correlations:

$$\log(1/K_i) = -11.502(\pm0.362) + 0.389(\pm0.118)LH \quad (12)$$
$n = 7, r = -0.82861, r^2 = 0.68659, s = 0.34737, F = 10.95333$
(2.1250%)

$$\log(1/K_i) = -6.302(\pm1.039) + 1.254(\pm0.337)DM \quad (13)$$
$n = 7, r = -0.85680, r^2 = 0.73411, s = 0.99732, F = 13.80456$
(1.3778%)

$$\log(1/K_i) = -16.368(\pm5.044) + 0.067(\pm0.033)VR_o + 1.252(\pm0.493)LH \quad (14)$$
$n = 7, r = 0.91838, r^2 = 0.84343, s = 0.58485, F = 9.69137$
(4.9070%), $r_c^2 = 0.2701$

$$\log(1/K_i) = -17.786(\pm2.953) + 0.111(\pm0.028)VR_i + 1.388(\pm0.286)LH \quad (15)$$
$n = 7, r = 0.96745, r^2 = 0.93597, s = 0.37402, F = 38.82284$
(0.7175%), $r_c^2 = 0.1117$

where VR_O and VR_I are the van der Waals volumes of the "out" and "in" substituents, respectively, being defined as "out" (out of the plane) and "in" (into the plane) when the molecules are viewed from the plane that approximately has the phosphate to the right and the carboxylate to the left, in analogy to the *si* face of PEP. The plot of equation 15 can be seen in Figure 5.

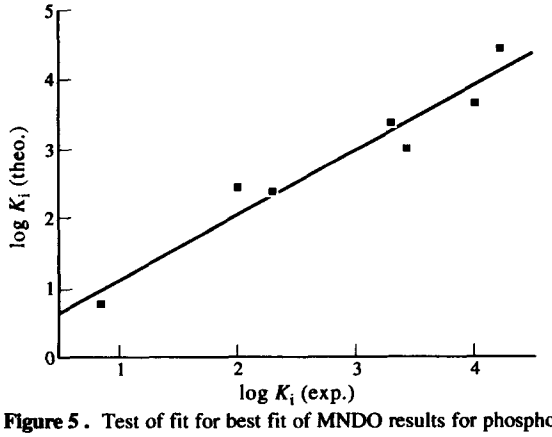


Figure 5. Test of fit for best fit of MNDO results for phosphoglycolate analogs.

From equations 12 to 15 it can be seen that, once again, dipole moment is one of the most important variables that describe the binding of these compounds to the active site. It is important to mention that the variable *LH* is here interpreted in terms of its relation to the electronegativity χ, which can be obtained approximately, following Mulliken's definition, as $\chi \equiv {}^1/_2$ (IP + EA) $\cong (E_{LUMO} - E_{HOMO})$. IP is the ionization potential and EA is the electron affinity. It seems that the increase in the ability of these compounds to 'pull' electrons, possibly during charge-transfer processes, results in a stronger binding force to the active site.

AM1

The results of the AM1 calculations for the PEP analogues produced the following best correlations:

$$\log(1/K_i) = -2.018(\pm0.205) + 0.289(\pm0.091)C_3R_{cis} \quad (16)$$
$n = 10, r = 0.74886, r^2 = 0.56079, s = 0.21152, F = 10.2148$
(1.2687%)

$$\log(1/K_i) = 0.138(\pm2.396) + 1.952(\pm0.500)C_3R_{cis} - 1.111(\pm0.509)C_cP \quad (17)$$
$n = 10, r = 0.85944, r^2 = 0.73863, s = 0.45183, F = 9.89103$
(0.9128%), $r_c^2 = 0.0001$

$$\log(1/K_i) = 15.417(\pm4.084) + 5.318(\pm1.382)qPO_4^{2-} + 4.345(\pm1.116)qR_{trans} - 1.804(\pm0.529)C_cP \quad (18)$$
$n = 10, r = 0.89978, r^2 = 0.80960, s = 0.41654, F = 8.50417$
(1.3976%)

where, additionally, C_cP is the interatomic distance between the carbonyl carbon and the phosphorus and qR_{trans} is the charge on the *trans* substituent. The colinearities of the variables of the last equation are: r_c^2 $(qPO_4^{2-}/qR_{trans}) = 0.2541$, r_c^2 $(qPO_4^{2-}/C_cP) = 0.20682$, and r_c^2 $(qR_{trans}/C_cP) = 0.0258$.

Equations 16 to 18 suggest that, as in the MM results, the bond distance between the double-bond end and the *cis* substituent is very important in determining the activity of PEP analogues. The difference in the contribution of the C_3R_{cis} variable to the activity between equations 16 and 2, about 6 times less in the former equation, can only be attributed to the different values of this interatomic distance calculated by both methods, which importantly still produce the same trend. Furthermore, the decrease in the distance between the phosphorus and carbonyl carbon atoms results in a better binding to the active site. Finally, active-site binding is also enhanced by means of specific electrostatic interactions with the phosphate and *trans* substituent of these compounds.

The results of the AM1 calculations for the phosphoglycolate analogues produced the following best correlations:

$$\log(1/K_i) = -12.087(\pm0.424) + 0.313(\pm0.138)LH \quad (19)$$
$n = 7, r = 0.71255, r^2 = 0.50773, s = 0.40706, F = 5.15706$
(7.2356%)

$$\log(1/K_i) = -5.619(\pm1.307) + 1.001(\pm0.424)DM \quad (20)$$
$n = 7, r = 0.72577, r^2 = 0.52675, s = 1.2544, F = 5.56518$
(6.4826 %)

$$\log(1/K_i) = -51.318(\pm 8.521) + 3.526(\pm 0.592)\,LH + 2.219(\pm 0.552)\,C_cP \quad (21)$$
$$n = 7,\ r = 0.94993,\ r^2 = 0.90236,\ s = 0.46184,\ F = 18.48418\ (0.9533\%),\ r_c^2 = 0.6386$$

The plot for equation 21 is shown in Figure 6.

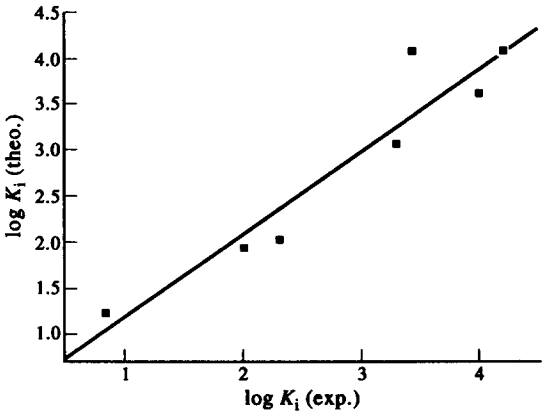


Figure 6. Test of fit for best fit of AM1 results for phosphoglycolate analogs.

The previous equations show that, the differences between the MNDO and AM1 results notwithstanding, dipole moment and the difference of energy LUMO–HOMO are the main factors that allow an adequate binding to the active site. Additionally, the interatomic distance between the phosphorus and the carbonyl carbon atoms seems to play an important role in this binding.

Discriminant analysis

The analysis of all the significant correlations obtained with the data from the previous methods revealed that the level of description of QSAR was far greatly improved when the PEPC inhibitors were separated into the PEP and phosphoglycolate analogue groups. This suggested that the separate descriptions of these two groups of compounds could be produced by the existence of two different mechanisms of interaction of these molecules with the active site. We proceeded to apply a two-variable graphical discriminant analysis[37,38] to the molecular properties of these compounds that were relevant in the previous multiple correlations. A clear separation of these two groups actually confirmed our previous hypothesis.

We were not able to find any clustering of the compounds into their belonging structural groups with the molecular properties obtained from the MM calculations. However, satisfactory clusterings were obtained with some pairs of molecular properties obtained with the MNDO and AM1 methods. With the MNDO properties, three cluster discriminants were obtained: qPO_4^{2-}/qCO_2^-, *DM/LH*, and qCO_2^-/LH, where qCO_2^- is the charge on the carboxylate group. The corresponding graphical display of the first discriminant is shown in Figure 7. With the AM1 properties, three cluster discriminants were also obtained: *DM/LH*, *POC/LH*, and qPO_4^{2-}/LH. The graphical display of the second discriminant is shown in Figure 8.

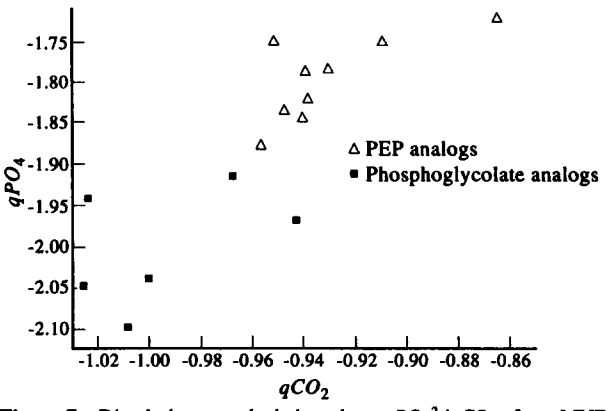


Figure 7. Discriminant analysis based on qPO_4^{2-}/qCO_2^- from MNDO calculations.

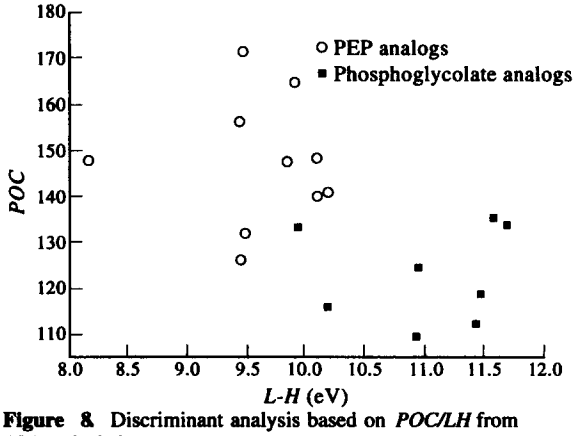


Figure 8. Discriminant analysis based on *POC/LH* from AM1 calculations.

Conclusions

No significant and regular correlations could be found with all PEPC inhibitors together (PEP and phosphoglycolate analogues). This fact, along with the cluster analysis, apparently suggests that these two groups of compounds can be separated by their own physical properties as an indication of actually different mechanisms of interaction with the active site. Net charges on the phosphate and carboxylate groups, along with other properties like dipole moment, the difference of energy LUMO–HOMO, and the angle P–O–C are responsible for the discrimination of these inhibitors into two groups that should behave differently when they interact with the active site of PEPC. Since it had never before been suggested that there could actually be more than one mechanism of interaction of these inhibitors with the active site of the enzyme, we should be cautious to emphasize that we are actually just referring to the initial binding of the inhibitors to the enzyme, and not to any subsequent reaction mechanism, if any.

In the case of PEP analogues, the interatomic distance between the double-bond end and the *cis* substituent seems to be the most important factor in determining the binding energy to the active site. It seems that a better spatial adjustment with the enzyme is achieved when this distance is increased, which may be due to the interaction

with a large pocket in the active site, since the van der Waals volume of this substituent was also significant in explaining the activity of these compounds. The results from the Hansch method also suggest that there is a large and hydrophobic pocket near to where the *cis* and *trans* substituents interact in the active site.

Another structural feature of the binding of these inhibitors with the active site is the fact that it is decreased when the distance between the phosphorus and the carbonyl carbon is also increased, as opposed to the case of the phosphoglycolate analogues, which has the opposite effect. This reveals that spatial adjustment of these ligands is important in determining the effectiveness of their binding, especially if we consider our suggestion of different interaction mechanisms for these two families of compounds.

Electrostatic features are also responsible for the different degrees of activity shown by these compounds. The increase in their dipole moments, the negative charge on the phosphate and the positive charges of the *cis* and/or *trans* substituents are important variables in explaining their ability to bind to the active site.

In the case of phosphoglycolate analogues, the increase in their ability to 'pull' electrons allows a stronger binding to the active site, probably by means of charge transfer interactions with basic regions in the enzyme. An increase in their dipole moments allows electrostatic interactions that enhance their active-site binding. A decrease in their non-polar saturated surfaces (which is made up from their substituents) increases their activity, which may be due to the interaction of their substituents with hydrophilic regions in the active site, adjacent to the hydrophobic regions with which the PEP analogue substituents interact.

It should be mentioned here that the properties of the C–C–O angle predicted by O'Leary in relation to the activity of some PEPC inhibitors were not confirmed by our studies, since neither the angles nor the trend she predicts were found. On the other hand, a significant correlation (not shown) was found with the charge on the phosphate ester oxygen and the P–O–C angle when some of the compounds were considered (phosphoglycolate, PEP phosphonate, S-PEP, phosphonopropionate, and phosphotioglycolate). Apparently, an increase in the negative charge of the above mentioned oxygen and in the P–O–C angle results in an increased binding to the active site.

The comparison of the correlations obtained by the three semiempirical methods described, shows some consistent results for MM and AM1 in the case of PEP analogues (i.e. the C_3R_{cis} variable) and for all methods in the case of phosphoglycolate analogues (i.e. the DM and LH variables). This might suggest that AM1 could be preferable for an initial survey of the geometric and electronic properties of compounds with the same sort of chemical groups (phosphocarboxylates).

We are currently performing *ab initio* calculations on all of these PEPC inhibitors to refine our correlations with better equilibrium geometries and electronic properties. This should also be very helpful for the calibration of the semiempirical methods used, in particular for the phosphate moiety. We are also trying to enhance our discriminant analyses to the PEPC allosteric activators, initially on the basis of PM3 calculations, in order to be able to explain the underlying physical properties that make a PEP analogue an inhibitor or an activator.

Experimental

The method of Hansch was applied after obtaining from the literature[15] all the relevant parameters (σ, π, E_S, MR) to describe the different substituents of the PEP analogues, and then a multiple regression analysis was performed to find the significant correlations.

MM[22] calculations were performed with the program PCMODEL, version 4.0 from Serena Software (Ergo Computing, Inc. for Lahey, 1990). Two compounds, S-PEP and phosphotioglycolate, were not treated due to the lack of parameters for the S–P bond. After a full geometry optimization of each molecule, internal rotations were allowed until the absolute minimum of energy was reached. Dipole moments, interatomic distances, bond angles and dihedral angles were obtained from these geometries. With the same software, we calculated for every molecule its van der Waals and solvent-accessible surfaces, along with their respective saturated non-polar, unsaturated non-polar and polar surfaces.

Quantum chemical calculations were performed with the program MOPAC, version 6.00 from Serena Software (Ergo Computing, Inc. for Lahey, 1990), implemented on several personal computers. Two different semiempirical methods were used: MNDO[23] and AM1,[24] with their standard parametrizations.[23–32] The initial geometries were those obtained from the previous MM calculations, and full geometry optimizations were performed separately with both methods. Once again, interatomic distances, bond angles and dihedral angles were obtained from the final equilibrium geometries. Several tests were done from different initial conformations to ensure the uniqueness of the equilibrium geometries found and their thermodynamic significance. The electronic properties that were obtained are dipole moment, ionization potential, heat of formation, orbital energies, net atomic charges, bond orders and HOMO and LUMO charge matrices.

It is known[32] that all semiempirical methods give too high heats of formation for pentavalent phosphorus compounds. This has been ascribed to the absence of d orbitals for P, which have an effect on the calculation of equilibrium geometries, as shown by *ab initio* calculations with moderately large basis sets. We could only hope, in this our initial study of the PEPC inhibitors, that we could still

see some trends that could be related to the biological activity of these compounds.

With the previously determined MNDO and AM1 geometries for each compound we proceeded to calculate the associated van der Waals surfaces and volumes with the program SAVOL, whose FORTRAN code was developed by R. S. Pearlman from the University of Texas at Austin. This program was compiled in a Silicon Graphics Iris INDIGO workstation. The atomic radii were those from the literature.[33,34]

Multiple regression and discriminant analysis, and graphical displays were done with standard personal computer software, including the SAS program from SAS Institute, Inc., through the assistance of IIMAS–UNAM.

QSAR studies that make use of quantum-chemical methods always face the problem of having too many variables among which a statistical screening is usually done to search for significant correlations. It has been reasonably suggested[36] that when too many variables are screened relative to the number of available observations, there is a risk of obtaining correlations fortuitously. In our study, a relatively large number of molecular properties were extracted from our calculations; however, an initial screening was done, based on what is known about the active site of PEPC. We selected only certain electronic and geometric properties in order to focus the statistical analyses to those sets of variables that could be relevant to the possible interaction with the known residues in the active site of the enzyme. It should be noted that the guideline of five compounds per parameter, frequently recommended for QSAR studies,[35] would require 10 compounds to obtain two-parameter equations and 15 for the three-parameter ones. Since experimental activity information is available for only six to 10 compounds in our study, we limited ourselves, basically for the sake of statistical significance, to up to two-parameter equations.

Acknowledgments

This work was supported by grants from the Dirección General de Asuntos del Personal Académico (DGAPA) of the Universidad Nacional Autónoma de México (UNAM). We would like to thank Dr Manuel F. Rubio-Arroyo for giving us a copy of the program SAVOL, Dr Ignacio Méndez from IIMAS–UNAM for allowing and helping us to work with the SAS program from SAS Institute Inc. at his institute, and Professor Silvia Alatorre-Frenk from Universidad Pedagógica Nacional for her advice on several statistical matters. We also thank Dr Rosario Muñoz-Clares and Dr Rogelio Rodríguez-Sotres for their critical contributions to this work. Finally, it is a pleasure to thank one of the reviewers of this paper for pointing out to us a valuable reference.

References

1. O'Leary, M. H. *Ann. Rev. Plant Physiol.* **1982**, *33*, 297.

2. Andreo, C. S.; González, D. H.; Iglesias, A. A. *FEBS Lett.* **1987**, *213*, 1.

3. González, D. H.; Iglesias, A. A.; Andreo, C. S. *Arch. Biochem. Biophys.* **1986**, *245*, 179.

4. Iglesias, A. A.; González, D. H.; Andreo, C. S. *Biochim. Biophys. Acta* **1984**, *788*, 41.

5. Stiborová, M.; Leblová, S. *Physiol. Vèg.* **1983**, *21*, 935.

6. Iglesias, A. A.; Andreo, C. S. *Biochim. Biophys. Acta* **1983**, *749*, 9.

7. González, D. H.; Andreo, C. S. *TIBS* **1989**, *14*, 24.

8. O'Leary, M. H. *Physiol. Vèg.* **1983**, *21*, 883.

9. Sikkema, K. D.; O'Leary, M. H. *Biochemistry* **1988**, *27*, 1342.

10. Liu, J.; Peliska, J. A.; O'Leary, M. H. *Arch. Biochem. Biophys.* **1990**, *277*, 143.

11. Rodríguez-Sotres, R. Doctoral Thesis, Facultad de Química, UNAM 1990.

12. Izui, K.; Matsuda, Y.; Kameshita, I.; Katsuki, H.; Wood, A. E. *J. Biochem.* **1983**, *94*, 1789.

13. Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1616.

14. Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175.

15. Hansch, C.; Leo, A. J. *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley; U.S.A., 1979.

16. Martin, Y. C. *J. Med. Chem.* **1981**, *24*, 229.

17. Miller, R. S.; Lane, M. D. *J. Biol. Chem.* **1968**, *22*, 6041.

18. Mukerji, S. K. *Arch. Biochem. Biophys.* **1977**, *182*, 352.

19. Wedding, R. T.; Rustin, P.; Meyer, C. R.; Black, M. K. *Plant Physiol.* **1988**, *88*, 976.

20. Rodríguez-Sotres, R.; Muñoz-Clares, R. A. *Arch. Biochem. Biophys.* **1990**, *276*, 180.

21. Janc, J. W.; O'Leary, M. H.; Cleland, W. W. *Biochemistry* **1992**, *31*, 6421.

22. Burkert, U.; Allinger, N. L. *Molecular Mechanics*, ACS Monograph 177, American Chemical Society; Washington, D.C., 1982.

23. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

24. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

25. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907.

26. Dewar, M. J. S.; Rzepa, H. S. *J. Am. Chem. Soc.* **1978**, *100*, 58.

27. Dewar, M. J. S.; McKee, M. L.; Rzepa, H. S. *J. Am. Chem. Soc.* **1978**, *100*, 3607.

28. Dewar, M. J. S.; Rzsepa, H. S. *J. Comp. Chem.* **1983**, *4*, 158.

29. Dewar, M. J. S.; Healy, E. *J. Comp. Chem.* **1983**, *4*, 542.

30. Dewar, M. J. S.; Reynolds, C. H. *J. Comp. Chem.* **1986**, *7*, 140.

31. Dewar, M. J. S.; Zoebisch, E. G. *J. Mol. Struct. (Theochem.)* **1988**, *180*, 1.

32. Dewar, M. J. S.; Jie, C. *J. Mol. Struct. (Theochem.)* **1989**, *187*, 1.

33. Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

34. Moriguchi, I.; Kanada, Y.; Komatsu, K. *Chem. Pharm. Bull.* **1976**, *24*, 1799.

35. Topliss, J. G.; Costello, R. J. *J. Med. Chem.* **1972**, *15*, 1066.

36. Topliss, J. G.; Edwards, R. P. *J. Med. Chem.* **1979**, *22*, 1238.

37. Dove, S.; Franke, R.; Mndshojan, O. L.; Schkuljev, W. A.; Chashakjan, L. W. *J. Med. Chem.* **1979**, *22*, 90.

38. McFarland, J. W.; Gans, D. J. *J. Med. Chem.* **1987**, *30*, 46.

(*Received in U.S.A.* 20 *June* 1994; *accepted* 5 *December* 1994)