Parabolic Relationships between Antibacterial Activity of Cephalosporins and β-Lactam Reactivity Predicted from Molecular Orbital Calculations[1]

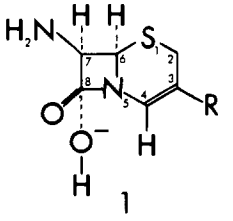
Donald B. Boyd,* David K. Herron, W. H. W. Lunn, and Wayne A. Spitzer

Contribution from the Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46206. *Received June 18, 1979*

Abstract: A molecular orbital index of reactivity of the β-lactam ring of model 3-substituted cephem structures is found to correlate in a rational manner with in vitro antibacterial activity of the corresponding cephalosporins. The index, called transition-state energy, is a measure of the ease of approach of a nucleophile toward the β-lactam ring in a mechanism patterned after that thought to occur in the inhibition of certain bacterial cell wall enzymes.

The β-lactam antibiotics, such as penicillins and cephalosporins, inhibit bacterial growth by differentially interfering with transpeptidase and carboxypeptidase enzymatic activities involved in biosynthesis of the peptidoglycan layer of bacterial cell walls.[2] The antibiotics do this by acylation of the enzymes. It follows that the reactivity of the β-lactam ring should be one of the factors determining antibacterial activity. Indeed, there is much evidence to support this statement.[3] A way to capitalize on this relationship is to use molecular orbital calculations to theoretically predict the reactivity of known and novel β-lactam structures. Such calculations, even if (*of necessity*) done semiempirically, should give for a related series of structures a measure of their *relative* reactivities. The purpose of this article is to describe correlations between predicted reactivities and observed antibacterial activities for some series of cephalosporins.

The theoretical index of reactivity which we have found to be most useful in understanding structure–activity relationships among cephalosporins is called the transition-state energy (TSE). It is a measure of how easy it is to form the initial complex of a model nucleophile, namely, OH^-, and a 3-cephem model structure with a substituent R at position 3. The structure of the "transition-state" complex **1** is the same as that



1

developed in earlier work.[4-6] The TSEs are computed by the CNDO/2 method[7] as the difference in total energies of **1** and the infinitely separated reactants (OH^- and the 7-NH_2-3-R-3-cephem). Because of approximations of the MO theory and lack of solvation effects,[4] the TSEs are not *equal* to activation energies, but one might expect them to be *roughly proportional* to them for a closely related series of compounds. CNDO/2 makes the complex appear to be about 130–140 kcal/mol more stable than the infinitely separated reactants, so that, the better the 3-R substituents can stabilize the impinging OH^-, then the more negative are the TSEs.

Biological activities are taken from several sources.[8] Antibacterial activity is expressed in terms of minimum inhibitory concentration (MIC) measured for 7-(2-thienylacetyl)cephalosporins against five Gram-negative, G(−), pathogenic microbes indicated in Figure 1. Because of the characteristically high variability[9] in assays of the MICs, each was measured more than once and an average taken. The averages for each of the five test organisms were then averaged together to obtain a general measure of G(−) activity.

Results are shown in Figures 1 and 2. In Figure 1 the data are for 7-NH_2-3-cephem structures with 3-CH_2R' substituents, where R′ ranges from a very poor leaving group (H) to very good leaving groups (heterocyclic thiols). In Figure 2 are plotted data for "direct" position 3 substituents (i.e., those without a methylene bridge). Also included in Figure 2 is 3-CH_3 because the hydride ion is such a poor leaving group that methyl could be considered as a direct substituent. In both figures, the 3-CH_3 data fall at the lower extreme of the −TSE scale and at the top extreme of the measured activity scale.

The data in Figure 2 can be considered to fall within a U-shaped curve. In Figure 1, there are few data points at very high −TSE to indicate whether this curve also bends up at high −TSE. However, the very fact of the scarcity of known 7-thiopheneacetylcephalosporins that have been synthesized and tested which have −TSE's falling at the upper end of the scale suggests that such structures have stability problems. Moreover, biological data for the 7-phenylglycylcephalosporins also lead us to believe that the biological activity should deteriorate at high −TSE. For instance, in the phenylglycyl series, the averaged G(−) MIC is very high and difficult to quantitate for the pyridiniummethyl side chain (because of stability problems[9]), but it is below 10 μg/mL for the tetrazoyl- and thiadiazolylthiomethyl side chains.[8] Thus, we feel that it is reasonable to assume a parabolic relationship for both figures and thereby obtain the following linear regressions. For the data in Figure 1

$$\mathrm{MIC} = (1.63 \pm 0.15)(\mathrm{TSE} + 136)^2 - 2.02$$

$$n = 13,\ r = 0.95,\ r^2 = 0.91,\ s = 6.08,\ P = 0.0000$$

For the data in Figure 2

$$\mathrm{MIC} = (0.80 \pm 0.27)(\mathrm{TSE} + 136)^2 + 3.25$$

$$n = 8,\ r = 0.77,\ r^2 = 0.59,\ s = 11.94,\ P = 0.0265$$

For the data in Figures 1 and 2 combined

$$\mathrm{MIC} = (1.02 \pm 0.19)(\mathrm{TSE} + 136)^2 + 2.25$$

$$n = 20,\ r = 0.78,\ r^2 = 0.61,\ s = 10.30,\ P = 0.0000$$

These regression equations give quite satisfactory correlations both in terms of the amount of variance accounted for (r^2) and in terms of the probability P based on F values that the relationships are fortuitous. The standard estimate of the error s is in units of μg/mL.

It would not be appropriate to take the journal space to present all conceivable or calculated regression equations based on the data in Figures 1 and 2, but a few comments are in order.

0002-7863/80/1502-1812$01.00/0

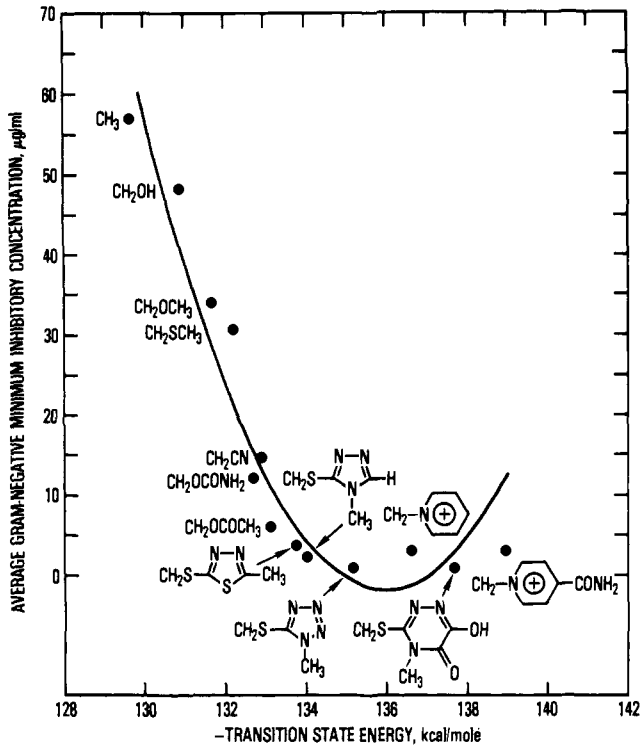


Figure 1. Average in vitro Gram-negative inhibitory concentration (in μg/mL) vs. −TSE. The MICs are from gradient plate assay for the thiophenacetyl cephalosporins with the 3-R groups as shown. An arithmetic average was taken of averaged MICs against N9 (*Shigella sonnei*), N10 (*Escherichia coli*), X26 (*Klebsiella pneumoniae*), X68 (*Enterobacter aerogenes*), and X514 (*Salmonella heidelberg*).

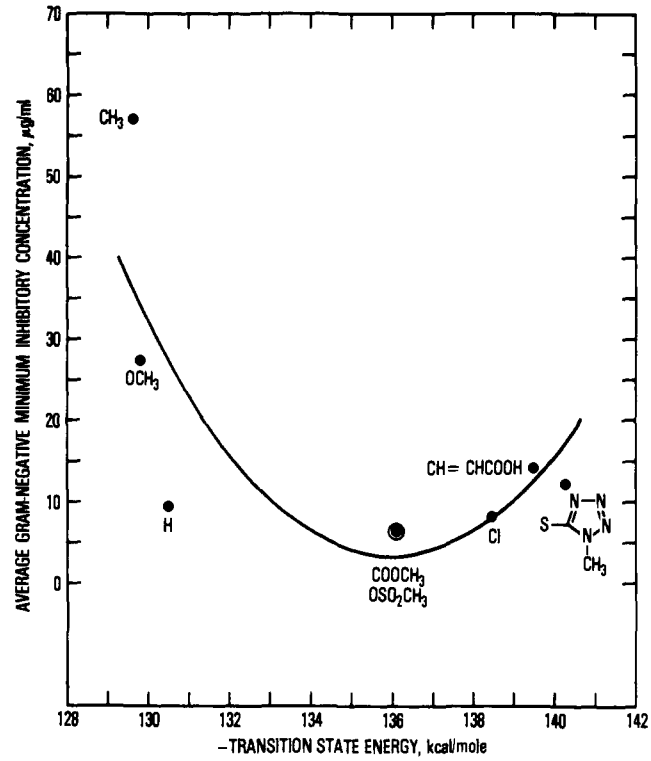


Figure 2. Average in vitro Gram-negative minimum inhibitory concentration (in μg/mL) vs. − TSE. The biological activities are expressed as described in Figure 1. The data for cephalosporins with 3-R = $COOCH_3$ and 3-R = OSO_2CH_3 fall on the same point.

The precision of the biological data should be kept in mind when interpreting the statistics.

The use of 136 in the above regressions is based on the fact that MICs are lowest for the cephem structures with −TSE's near this value. In other words, 136 is taken to be the minimum of the parabolas. Depending on the sample of data taken from Figures 1 and 2, the correlation coefficient r can be improved by up to 0.02 by refining the position of the minimum more exactly between 136 and 137 kcal/mol.

It could be argued that inclusion of the pyridinium models in Figure 1 is not valid because these are charged side chains and require a counterion[6] to make the magnitude of the TSEs comparable to those for the remaining side chains. The value of the TSE for the pyridinium models is very sensitive to the spatial location of the counterion (Cl^-). Leaving the two pyridinium models out of the regression for Figure 1 improves r^2 to 0.93 and lowers s to 5.77 μg/mL.

Using a regression equation of the form MIC = aTSE^2 + bTSE + c gives r^2 = 0.77 for the 13 points in Figure 1, r^2 = 0.58 for the 8 points in Figure 2, and r^2 = 0.60 for all 20 data points. All three predicted curves are parabolic in shape.

The fit of the data to the plotted regression equation is obviously better in Figure 1 than in Figure 2. Of course, there are various biological factors which would affect how good a fit is achievable. These include (1) penetration of the antibiotic through the outer bacterial membrane to the sites of action, (2) reversible binding affinity of the antibiotic with the target enzymes, (3) rate of covalent attachment of the inhibitor molecules to the enzymes, and (4) rate of deacylation of the acyl-enzyme intermediate. These quantities have not been measured for the set of compounds treated here. Then, too, it is possible that the CNDO/2 MO method may not reflect the electronic effects of the drastically different direct 3-R groups as well as in the case of the 3-CH_2R' substituents. For instance, it is known that the usual CNDO/2 parameters[7] do not account well for the polarity of the C-H bonds.[10] One proposal to improve the charge distributions of methyl vs. hydrogen has been to use −5.03 eV instead of −7.176 eV for the $(I_s + A_s)/2$ parameter of hydrogen.[10] With standard parameters the 3-H atom has a CNDO/2D net atomic charge of +0.0352, and 3-CH_3 has a group charge of −0.0013. With the new parameter, the CNDO/2D charge on 3-H is +0.1286 and that on 3-CH_3 is −0.0075. Thus, in the ground-state cephem models the methyl substituent is calculated to have a net electron-withdrawing effect with respect to 3-H, and the amount of withdrawal is relatively greater with the new parameter. The −TSEs of the 3-CH_3 and 3-H models differ by only 0.9 kcal/mol when standard parameters are used (Figure 2), but differ by 2.25 kcal/mol with the new parameter. A greater gap like this might put the 3-H data point in better relation to the 3-CH_3 data point, although the TSE scale changes significantly with the alternate CNDO/2 parameterization.

The relationship between the computer-generated quantity (TSE) and the microbiological assay is intuitively understandable as follows. A small magnitude for TSE corresponds to a less reactive β-lactam ring. Hence biological activity may be expected to be poor (MICs high) because the compound (which the calculations attempt to model) is relatively poor at acylating the bacterial cell wall enzymes. On the other hand, a large magnitude for TSE corresponds to a very reactive β-lactam, and hence the compound may have stability problems. Biological activity would then be reduced because of reaction before reaching the appropriate receptor sites. Thus, the parabolic relationship results.[11]

The correlations we find extend earlier ideas[3] that chemical reactivity and biological activity should be related. Despite the multitude of factors[2-4,9] (some of which were alluded to above) which are involved in eliciting a biological response, our results suggest that, for a particular series of related cephalosporins tested against particular organisms, the reactivity of the β-lactam ring may be controlling relative activities.

Among the electronic effects of the 3-substituents influencing the TSE and MIC values are the following. One is

the inductive effect of the 3-R group. The degree of electron withdrawal, especially via the C_3 p_π orbital, influences both the electrophilicity of the β-lactam carbonyl carbon and the stability of the transition-state structure involved in opening the β-lactam ring. For instance, the methyl group withdraws σ electrons, but donates π electrons to the cephem ring. The latter donation makes the TSE and MIC less favorable for the 3-CH_3 structure compared to the 3-H structure. The 3-H substituent, of course, does not donate π electrons. The other effect is leavability[4,5] (nucleofugality[12]) because the ability of the R′ group of the 3-CH_2R' side chain to depart with a bonding pair of electrons will influence the ease of opening of the β-lactam ring at its site of action in the cell wall enzymes. The TSEs reflect how well the 3-substituent stabilizes the extra electronic charge introduced into the cephem system by our model nucleophile OH^-.[4,5] The TSEs are thought to be related to both the leavability of the R′ group[13] and the π electron-accepting ability of the side chain.

The MICs in Figure 2 drop rapidly to a low plateau as −TSE starts to increase from the 3-methyl value. In Figure 1, the drop is less steep at low −TSE because the inductive effect is screened by the intervening methylene moiety of the side chain. At intermediate TSE, leavability of the R′ moiety of the 3-CH_2R' side chains in Figure 1 more than compensates for any attenuation of the inductive effect and leads to G(−) MICs lower than in Figure 2. By overlaying Figures 1 and 2, it can be seen that the two regression curves cross in the vicinity of the CH_2CN/CH_2OCONH_2 and Cl data points. Thus, the contributions from inductive and leaving-group abilities of the 3-R substituents can combine to result in structures which fit either curve. While it may be argued that the results here do not alone justify the classification of leaving group vs. direct substituents, our findings are at least consistent with other evidence[5,14] that the leaving-group mechanism[4] is relevant to the chemistry and biology of cephalosporins.

Acknowledgments. We thank M. M. Marsh, P. Roffey, and C. Jochum for helpful suggestions on the manuscript.

References and Notes

(1) Part 11 of the series "Electronic Structures of Cephalosporins and Penicillins". See ref 5 for part 9.
(2) Yocum, R. R.; Waxman, D. J.; Rasmussen, J. R.; Strominger, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 2730. Ghuysen, J.-M.; Frère, J.-M.; Leyh-Bouille, M.; Coyette, J.; Dusart, J.; Nguyen-Disteche, M. *Annu. Rev. Biochem.* **1979**, *48*, 73. Boyd, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5239; *J. Med. Chem.* **1979**, *22*, 533, and references cited therein.
(3) See, e.g., Boyd, D. B. *J. Med. Chem.* **1973**, *16*, 1195. Yamana, T.; Tsuji, A. *J. Pharm. Sci.* **1976**, *65*, 1563. Indelicato, J. M.; Dinner, A.; Peters, L. R.; Wilham, W. L. *J. Med. Chem.* **1977**, *20*, 961.
(4) Boyd, D. B.; Hermann, R. B.; Presti, D. E.; Marsh, M. M. *J. Med. Chem.* **1975**, *18*, 408.
(5) Boyd, D. B.; Lunn, W. H. W. *J. Med. Chem.* **1979**, *22*, 778.
(6) The geometry fo the 7-NH_2-3-R-3-cephem is held fixed upon addition of OH^- to form **1**. The OH^- is 1.50 Å away from the α face of C_8 and oriented perpendicular to the C_8–C_7–C_6 plane. Other structural details are given in ref 4. Uncertainty in the TSEs due to certain geometrical assumptions about the side chain at position 3 is often 1% or less, but can be as high as 4%. For instance, the −TSE's for the pyridiniummethyl models must be regarded as having a relatively large uncertainty with respect to the neutral side chains because they had to be calculated with a Cl^- counterion arbitrarily placed with respect to the pyridinium ring at the optimized position referred to by Allinger, N. L.; Kao, J.; Chang, H.-M.; Boyd, D. B. *Tetrahedron* **1976**, *32*, 2867. The CNDO/2 optimized position is quite different than that expected by other methods (Jordan, F. *J. Am. Chem. Soc.* **1975**, *97*, 3330). Replacement of the 4-COOH and the *N*-acyl group of cephalosporins by hydrogens in **1** is a simplifying approximation justified on the basis of roughly constant, additive effects that substituents were found to have for a given nucleus structure. Some of the models treated herein, despite the simplifications, had over 100 valence atomic orbitals and took over 2 h IBM 370/158 CPU time per TSE. Although treatment of the models by some ab initio method is, in principle, preferred, such calculations would be impractical considering the number and size of molecules to be treated. Moreover, just as the goodness of semiempirical MO results depends upon cancellation of errors (from underlying approximations and assumptions) within a related series of structures, the predictions of ab initio methods are subject to cancellations of errors from additional factors (basis set, correlation). For further expounding on the choice of method, see Dewar, M. J. S.; Haddon, R. C.; Li, W.-K.; Thiel, W.; Weiner, P. K. *J. Am. Chem. Soc.* **1975**, *97*, 4540.
(7) Pople, J. A.; Beveridge, D. L. "Approximate Molecular Orbital Theory"; McGraw-Hill: New York, 1970. An spd basis set is used along with standard parameters.
(8) Lunn, W. H. W., unpublished data. Spitzer, W. A., unpublished data. Chauvette, R. R.; Pennington, P. A. *J. Am. Chem. Soc.* **1974**, *96*, 4986. Webber, J. A.; Ott, J. L.; Vasileff, R. T. *J. Med. Chem.* **1975**, *18*, 986. Webber, J. A.; Ott, J. L. In "Structure-Activity Relationships among the Semisynthetic Antibiotics", Perlman, D., Ed.; Academic Press: New York, 1977; p 161.
(9) Gorman, M.; Ryan, R. W. In "Cephalosporins and Penicillins: Chemistry and Biology", Flynn, E. H., Ed.; Academic Press: New York, 1972; p 532. Initial raw MIC data commonly have an uncertainty of ±50%, and it is sometimes almost as high as ±100%. By extensive retesting and averaging the uncertainty in the average can be brought down to a reasonable level (ca. 20%).
(10) Hermann, R. B. *J. Am. Chem. Soc.* **1970**, *92*, 5298.
(11) It may be that for a particular series of 7-acylcephalosporins (thiophenacetyl in our case) a U-shaped curve could be drawn when chemical reactivity is plotted against in vitro G(−) activity. The curve would define the activity which could usually be expected from the electronic contributions of the side chain at position 3. However, there may be a number of outliers not fitting the curve because of various possibilities: e.g., the side chain itself may undergo some reaction before the antibiotic reaches the receptor site, the compound may not be pursued to purity as high as obtained for the compounds in our series, a leaving group may be involved that has its own toxicity (see ref 5), or the transport properties across the outer membrane may be affected by the side chain. Thus, there is potential for gaining additional information from the outliers.
(12) Stirling, C. J. M. *Acc. Chem. Res.* **1979**, *12*, 198.
(13) Boyd, D. B., unpublished data. The magnitude of CNDO/2 reaction energies for breaking H–SR″ bonds, where R″ are heterocyclic thiols, is inversely proportional to the magnitude of the TSEs for the 7-NH_2-3-CH_2SR''-3-cephems. Hence the TSEs are related to theoretical gas-phase acidities for closely related structures.
(14) Lunn, W. H. W., published data.