Correlation of biological activity in β-lactam antibiotics with Woodward and Cohen structural parameters—a Cambridge database study

Ashwini Nangia,* Kumar Biradha and Gautam R. Desiraju *

School of Chemistry, University of Hyderabad, Hyderabad 500 046, India

Biological activity in β-lactam antibiotics has been correlated previously with one of two geometrical parameters, the Woodward height-of-pyramid, *h*, or the Cohen lactam O-atom to carboxylate C-atom distance, *c*. Here we describe a joint correlation of the two empirical parameters, *h* and *c*, with biological activity and the implications of such a correlation to enzyme–drug interactions in the β-lactam class of antibiotics. The Cambridge Structural Database (CSD) was the natural method of analysis and a substantial number of β-lactam structures (114) has been examined. It has been found that molecules with either high *h* and high *c* values or low *h* and low *c* values correspond to active skeletons. In contrast, molecules with high *h* and low *c* or low *h* and high *c* values are largely devoid of activity. The behaviour of outliers in both sets of populations has been explained on the basis of chemical and stereochemical factors. This correlation is far superior to the correlation of biological activity based either on *h* or *c* alone. Such an observation, namely that activity is a multivariate phenomenon, hints that a third parameter, which could be a torsional parameter about a non-bonded vector, is probably a better indicator of activity. Using crystal structure information on the binding of β-lactam antibiotics to their target enzymes, transpeptidases and β-lactamases, this parameter is shown to be an overall shape factor. The joint correlation of *h* and *c* parameters with biological activity provides a better understanding of structure–activity relationships and should find application in the rational discovery of new β-lactam antibiotics.

Introduction

The antimicrobial activity of β-lactam antibiotics is attributed to covalent bonding of the amide carbonyl group with the active site serine residue of enzymes, which are termed penicillin-binding proteins (PBP). Bicyclic β-lactam antibiotics such as penicillins, cephalosporins and thienamycins show wide-ranging therapeutic activity and have been in clinical use for the treatment of infectious diseases.[1-4] Their therapeutic efficacy is derived from their ability to disrupt bacterial cell-wall synthesis by inhibiting transpeptidase enzymes which catalyse the cross-linking reaction of D-alanyl peptides on peptidoglycan strands of the growing cell-wall.

The active site region of PBPs is a highly conserved tetrad of Ser–X–X–Lys residues. Structurally and evolutionarily-related penicillin recognizing enzymes, such as carboxypeptidases, transpeptidases and β-lactamases, have been co-crystallized with antibiotics and the role of amino acid residues in the formation of the initial non-covalent complex and acyl-enzyme intermediate studied with X-ray crystallography.[5-8] It is suggested that the hexapeptide Val–X–Ser–X–X–Lys is a model for PBP from *Streptomyces* R61 based on the 2.8 Å resolution X-ray structure of the protein.[5]

In their search for drugs with better therapeutic properties, greater selectivity, increased potency and stability to resistant bacterial strains[9,10] medicinal chemists have, over the years, made a large number of modifications of the active β-lactam core. The synthesis, evaluation and clinical trials with a large number of penicillins, cephalosporins, penems, thienamycins, olivanic acids, clavulanates, nocardicins and sulfazecins has, to date, produced more than 150 antibiotics. In spite of this intense and continuing interest in these compounds, much remains to be learned about the fundamental factors which influence their antibacterial activity and about the structural and conformational requirements for improved activity.

The biological activity of β-lactam antibiotics has been empirically correlated with several geometrical parameters. The most well-known of these is the Woodward parameter, that is the height of the pyramid (*h*) of the N-atom from the plane containing the three adjacent carbon atoms on the bicyclic β-lactam skeleton.[11,12] This is illustrated schematically in Fig. 1. This parameter is important because it reflects, through amide resonance, the strength of the C–N bond which must be of intermediate value for optimal antibacterial activity (*h* ~ 0.25–0.50 Å). If the C–N bond is too strong (*h* < 0.05 Å), it will be unreactive and covalent inactivation of serine peptidase will not occur, whereas if it is too weak, the antibiotic may be destroyed by reaction with random nucleophiles before it can reach the site of action. A geometrically equivalent parameter which has also been widely used[12,13] is the sum of bond angles at the N-atom (ΣN). In effect, *h* and ΣN are measures of the deviation from planarity in the amide portion of the antibiotic molecule.

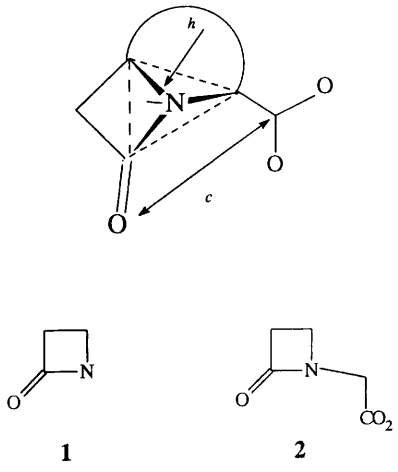


Fig. 1 Schematic representation of the Woodward height-of-pyramid parameter, *h*, and the Cohen distance parameter, *c* (top) and sub-structures **1** and **2** used in the CSD searches (bottom)

Non-fused monocyclic β-lactams, for example, have $h \sim 0$ Å ($\Sigma N \sim 360°$) and are usually inactive because they have a very strong C–N bond. A more direct measure of amide resonance inhibition and C–N bond strength is the C–N bond distance (r) and this too has been correlated with biological activity in a large number of structures.[13] Because of amide resonance, r is chemically correlated with h. In summary, h, ΣN and r are equivalent parameters indicating the extent of non-planarity at the N-atom and so a consideration of more than one of these three parameters is superfluous providing little additional information about structure–activity relationships. Accordingly, we will confine our further discussion to h.

A geometrically and chemically independent structural parameter which is related to the biological activity of β-lactam antibiotics was defined by Cohen[14] who studied conformational changes in the thiazolidine ring of penicillins and their effect on activity. It was shown that a pseudo-equatorial orientation of the acidic carboxyl group, wherein it is nearer to the β-lactam carbonyl O-atom, is decisive in promoting antibiotic activity. A three-dimensional conformational analysis in nine penicillins, cephalosporins and penems showed that the distance, c, of the lactam carbonyl O-atom to the carboxylate C-atom lies in the range 3.0–3.9 Å for six active structures, as compared to a higher range of 4.1–4.3 Å for the three inactive compounds. This value c is referred to as the Cohen parameter (Fig. 1). Thus, Woodward's h parameter ascribes biological activity to the chemical reactivity of the amide bond, whereas Cohen's c parameter lays emphasis on the 3-dimensional conformation of the antibiotic molecule.[15]

There have been numerous attempts in correlating the Woodward parameter, h, or the Cohen parameter, c, with activity in the search for potent and effective antibiotics. Although these investigations have contributed to our understanding of the factors that influence biological activity in β-lactams, they are somewhat limited in their scope and predictability. Firstly, all such studies have used either one or the other of the two independent parameters, h and c. Secondly, the number of examples considered to propose the utility and test the validity of these parameters has always been limited to a small and chemically homogeneous group.

Given the importance of three-dimensional structure and conformation in determining the activity of β-lactam antibiotics, the Cambridge Structural Database (CSD) appeared to us to be a convenient and reliable storehouse for this geometrical information[16,17] The utility of small-molecule crystallography and the CSD in elucidating drug–receptor interactions has been clearly established.[18,19] However, CSD studies specific to β-lactam antibiotics have not been reported. A preliminary search of the CSD (Version 5.1) showed that around 150 entries contain the β-lactam skeleton, **1** (Fig. 1). An overwhelming majority of these compounds appeared to have originated from pharmaceutical studies. It was felt therefore that a CSD analysis of these crystal structures would provide a range of h and c values for a large number of diverse antibiotics. We sought to examine possible relationships between the seemingly independent parameters h and c especially in connection with the following questions: (*a*) What particular structural features lead to high activity and potency? (*b*) Why are many molecules with 'favourable' h and c values devoid of antibacterial action? (*c*) Is a molecule inactive because of poor recognition by PBP, or because of inherently low chemical reactivity and possibly a different mechanism of action? (*d*) Is it possible to rationally modify an inactive structure and make it active? Since about two-thirds of all antibiotics today belong to the β-lactam category, an answer to these and related issues is expected to be beneficial to the medicinal chemist because it will: (*a*) provide a deeper insight into the biochemical mode of action, (*b*) facilitate in the rapid elimination of possible false leads and (*c*) lead to the systematic modification of functional groups in the non-β-lactam part of the antibiotic. Accordingly, the efficiency of the drug design cycle for new β-lactam antibiotics will be accelerated.

Results and discussion

Database results

Geometrical details from a total of 114 β-lactams containing fragment **2** (Fig. 1) were retrieved from the CSD. Fig. 2, which is a histogram of the Woodward h values in the range 0.00–0.60 Å, shows three regions of preference. The monocyclic β-lactams (monobactams) are the least populated category with h in the range 0.05–0.10 Å. In compounds with intermediate h values (0.20–0.25 Å), the β-lactam ring is fused to a six-membered ring, for example cephalosporins. A 'no-mans land' around 0.30 Å is followed by a highly populated region with h values between 0.40–0.50 Å corresponding to the penicillin class of antibiotics. There are very few antibiotics with h between 0.50–0.60 Å (highly pyramidal N-atom) and none with $h > 0.60$ Å. The region between 0.50–0.60 Å includes carbapenems and clavulanates. The latter are potent β-lactamase inhibitors,[20] which are themselves devoid of any antimicrobial activity, but are synergists in combination with β-lactam antibiotics. A visual inspection of Fig. 2 shows that penicillins have been studied more extensively by crystallographers than cephalosporins and monobactams and the preponderance of pencillin-like structures in the histogram may well reflect the fact that they are the oldest class of antibacterial agents (pre-1940s). The emergence of cephalosporins (1950s), cephamycins, penems, carbapenems, oxacephems, clavulanates (1970s) and monobactams (1980s) in later years as chemotherapeutic agents[3,4] is reflected in the fewer number of crystal structures for these skeletons in the database.

The histogram of the Cohen distance (c) given in Fig. 3 reveals that the number of structures with c between 3.0–3.9 Å (active range as defined by Cohen) is less than half the total structures retrieved. More than 60 structures have c in the range 4.0–4.5 Å with local maxima around 4.0 and 4.4 Å. While these values correspond to the inactive region as defined by Cohen, many of these compounds are clearly active, such as penicillin G and clavulanic acid. Inasmuch as complications such as pseudorotation would cloud such issues, it is likely that this limitation in Cohen's classification arose from the fact that too small a number (9) of known active and inactive compounds was studied. This highlights, in general, the problem with using limited and specialized data samplings in statistical studies of structure–activity relationships. The ever-increasing size of the CSD, which now contains more than 120,000 mostly accurate crystal structures, makes it the method of choice for analysis whenever geometrical parameters of small-molecules are sought to be related to chemical or biochemical activity.[21]

It may be noted that the h parameter is characteristic of the molecular structure while the c parameter is related to conformational properties. For example, penicillin and its 3-epimeric carboxylate will have similar h values, but very different c values.[22] These parameters are therefore independent and the next step in the analysis was to examine the h versus c scatterplot for the 114 hits, see Fig. 4, which combines the information in Figs. 2 and 3. There is a more or less uniform distribution of points in the h range 0.00–0.60 Å and the c range 3.0–4.5 Å, but other than expanding the range of c to 4.5 Å (which is also seen in Fig. 3), this scatterplot is uninformative.

We reasoned at this stage that an h–c scatterplot for only the active β-lactam skeletons should be more informative. Accordingly, eight bicyclic β-lactam cores were selected—penams, cephems, clavams, penems, carbapenems, oxapenems, carbacephems and oxacephems (Fig. 5). The identification of these skeletons as active cores was based on a perusal of leading

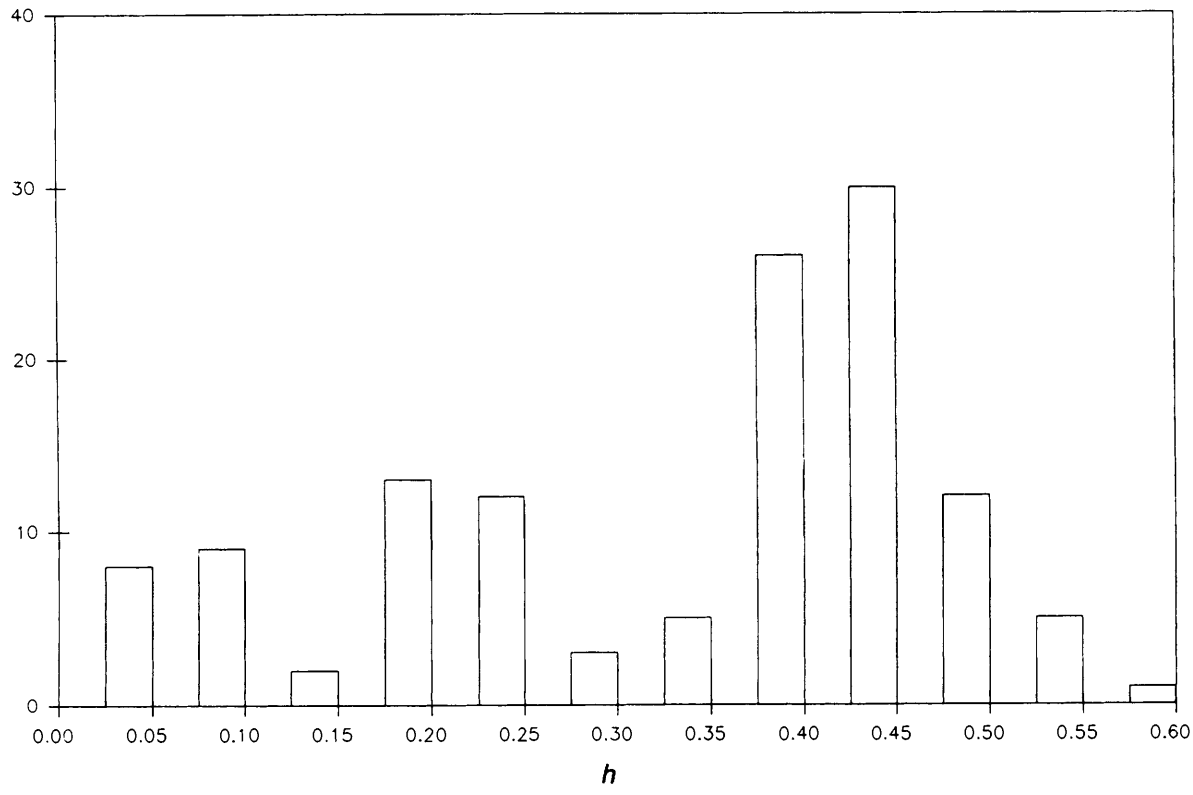


Fig. 2 Histogram of the Woodward *h* values for the 114 structures in the present study

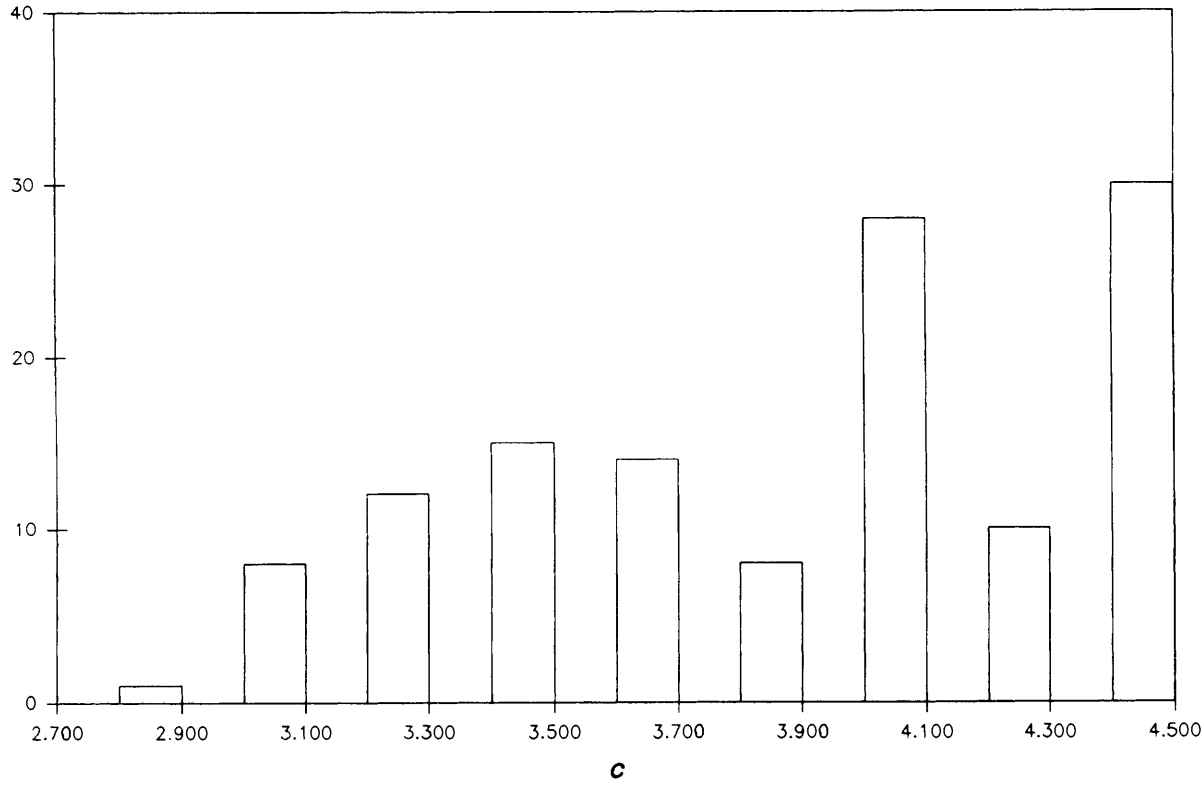


Fig. 3 Histogram of the Cohen *c* values for the 114 structures in the present study

reviews on β-lactam antibiotics[1-4] and the knowledge that each of these groups contains a commercial drug or at least an advanced clinical candidate. A total of 80 hits (out of the original 114) was obtained for these active skeletons and their *h*–*c* scatterplot is given in Fig. 6. A comparison of Figs. 4 and 6 is striking because it shows that the active skeletons lie in a limited region of the *h*–*c* space, mostly along the diagonal of the scatterplot. Active compounds are characterized by a combination of either high *h* (0.35–0.50 Å) and high *c* (3.5–4.5 Å) values or by low *h* (0.15–0.25 Å) and low *c* (3.1–3.6 Å) values. The off-diagonal possibilities, namely high *h* and low *c* or low *h* and high *c* correspond to the inactive compounds (Fig. 7).

Once the Woodward and Cohen parameters are taken together, a far more rigorous interpretation is obtained about biological activity than is possible by analysis with either parameter alone. This is illustrated by a few representative

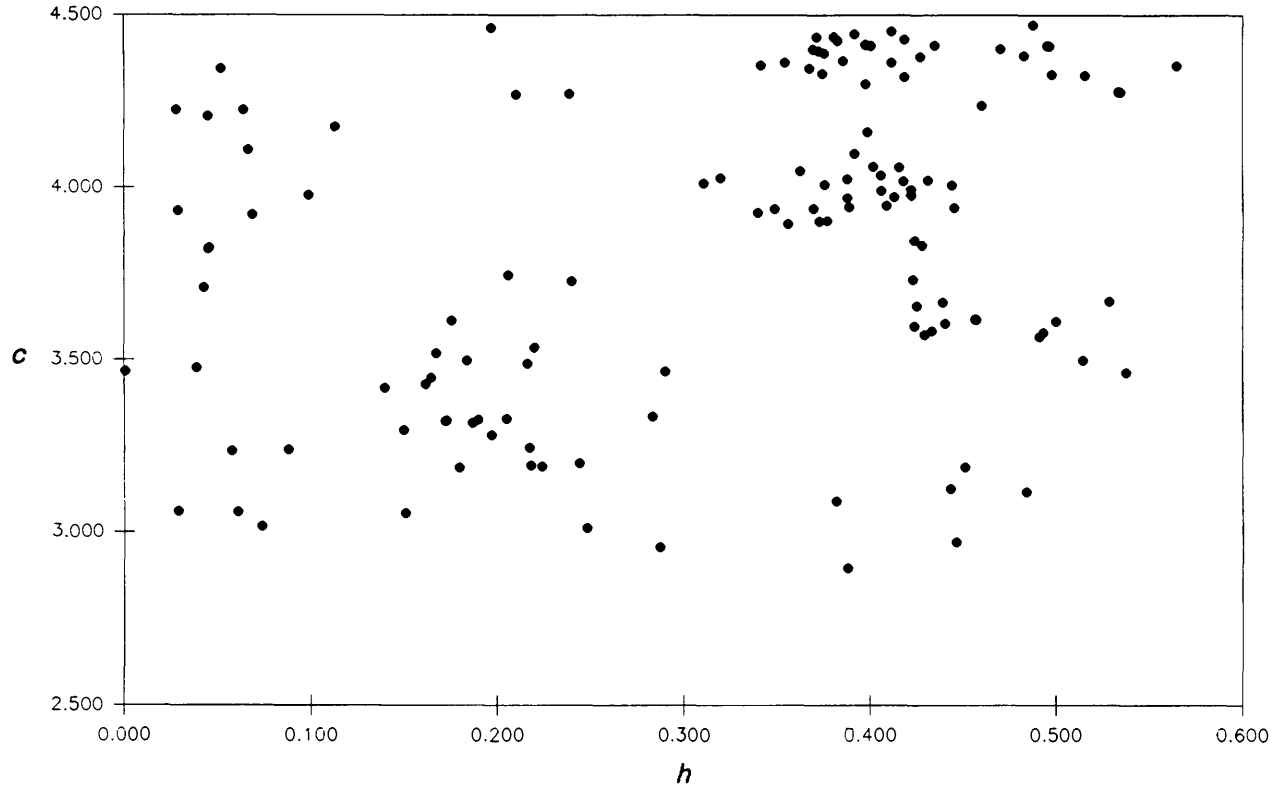


Fig. 4 Scatterplot of *h versus c* values for all 114 β-lactams in this study

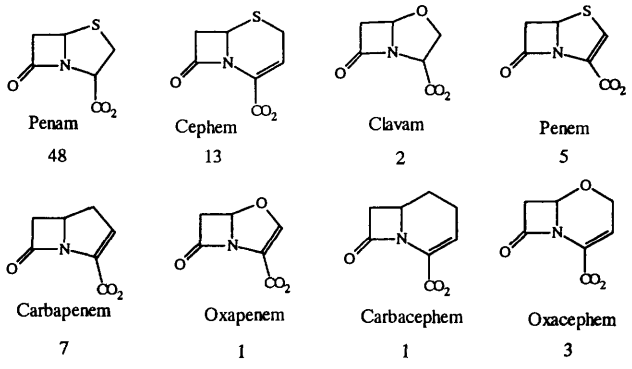


Fig. 5 Active skeletons considered in Fig. 6. The numbers below the structures refer to the populations of each group in the plot.

examples. The CSD REFCODE, *h* and *c* values of the 3α- and 3β-penams **3** and **4**,[23,24] the Δ^2-cephem **5**,[13] the 4α- and 4β-carbacephams **6** and **7**[25–27] and the Δ^1-carbapenem **8**[28] are displayed in Fig. 8. Predictions based on the *h* parameter alone should lead one to conclude that epimeric penams **3**, **4** should be active because of their folded skeletons and reactive amide carbonyl group, while Δ^2-cephem **5** should be predicted to be inactive because of its near-flat shape. In such an analysis, the stereochemistry (α or β orientation) of the $-CO_2H$ group, which contributes significantly to bioactivity (through its intermolecular interactions), is completely ignored. Predictions based on the Cohen parameter alone could also be equally misleading. Benzylpenicillin **3** with a *c* distance of 4.43 Å is predicted to be inactive on this count, whereas it is actually one of the earliest-discovered antibiotics. The *h*–*c* scatterplot in the present study resolves this ambiguity with the working hypothesis that all active leads should have either high *h* and high *c* distances or low *h* and low *c* distances. That penicillin G **3** must be a good serine peptidase inhibitor because it falls along the 'active' diagonal is trivially obvious. More importantly, the inactivity of epimeric 3β-penam **4** is now revealed. Though its *h* value (0.37 Å) is favourable for activity, the *endo* carboxylate group forces a smaller value for the *c* parameter (3.14 Å) and hence this compound is an off-diagonal point in Fig. 6. The reverse behaviour is observed in the homologous series of carbacephams **6** and **7**. It is known that cepham analogs in which the 4-carboxy group is oriented β (*endo* face) are about 8–10 times more active than their 4α (*exo* face) counterparts,[26,29] although the skeletons are clinically not very important because the absolute MIC values are high. Once again, the more active β-carboxy cephem **7** has low *h* and low *c* distances, whereas the inactive 4α-epimer **6** lies in the off-diagonal region (low *h* and high *c*). The well-known inactivity of Δ^2-cephem **5** is further reinforced because not only does it have a very low *h* but also a very high *c* distance. Finally, Δ^1-carbapenem **8** whose skeleton is more folded than penicillins and penems is found to be completely devoid of antibacterial activity.[28] Here, our two-parameter analysis also fails to correctly predict the bioprofile of this compound which occurs in the high *h* and high *c* region of the 'active' diagonal in Fig. 6. While we have demonstrated the superiority of simultaneously employing the two independent structural parameters, *h* and *c*, for predicting biomolecular properties, an inherent limitation of any such geometry-based approach is that a predefined structure and conformation of a potential bioactive molecule are necessary but not sufficient conditions for antibacterial activity. There are other factors which are responsible for the 'perfect fit and reactivity' of a substrate in the enzyme pocket.

A closer examination of Figs. 6 and 9(*a*) shows three outlier sets, the geometrical details of which are now delineated: (*a*) the bicyclo[3.2.0] β-lactam **4** synthesized by Hanessian *et al.*[24] is an outlier because the carboxy group is forced into the *endo* orientation by synthetic design. Consequently, *c* is very small (2.90 Å) whereas the *h* value (0.39 Å) is normal for a β-lactam fused to a five-membered ring. There are a few other examples of bicyclic molecules (**9**–**11**) in which the carboxylate group is in the unnatural β stereochemistry.[30–32] These three structures belong to the carbapenam or carbacepham skeleton and appear randomly in the 'inactive' *h*–*c* scatterplot (Fig. 7). (*b*) Clavulanate **12** is not part of the cluster.[33] This molecule has an unusually high *h* (0.57 Å) because the clavam O-atom

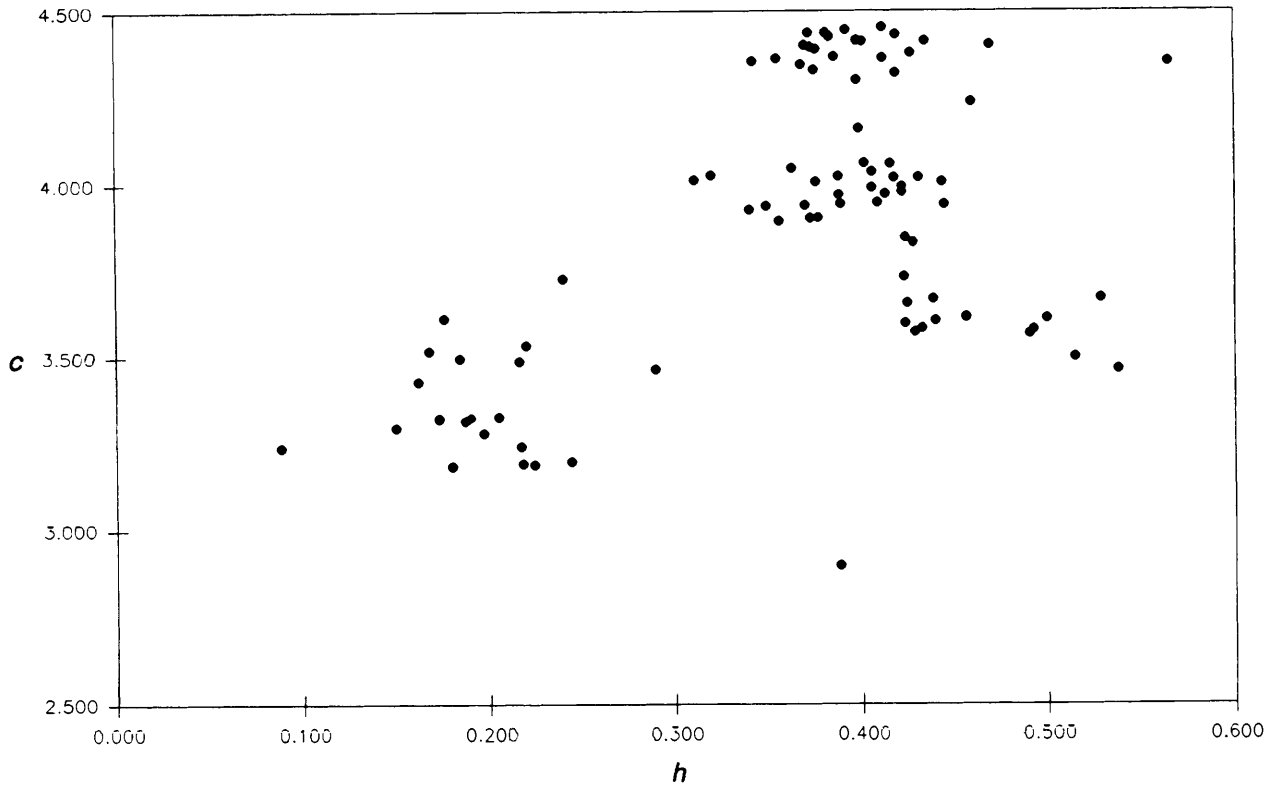


Fig. 6 Scatterplot of *h versus c* values for the 80 β-lactams with active skeletons. Note that the points cluster along the diagonal. A few outliers in the high *h* and moderate *c* region may be observed.

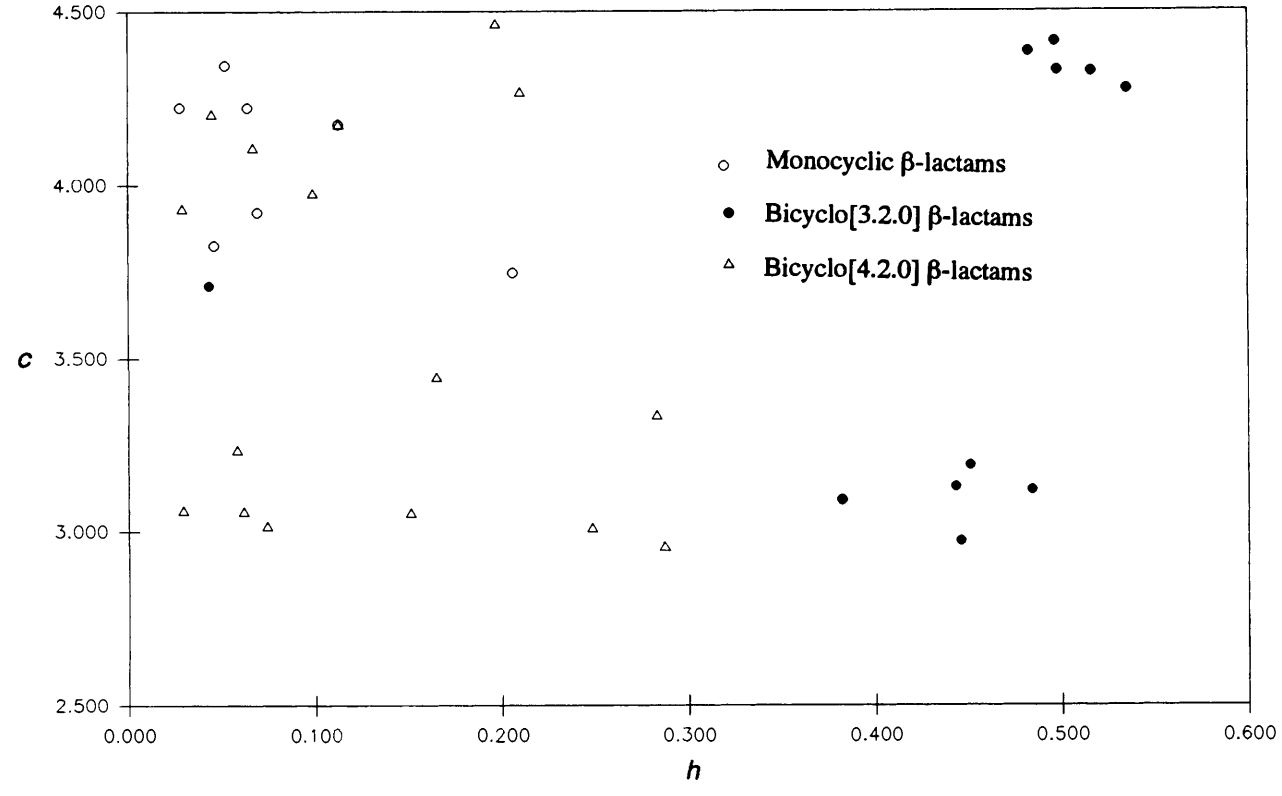


Fig. 7 Scatterplot of *h versus c* values for the 30 inactive β-lactams. Notice the wide scatter of points.

inductive effect strongly inhibits amide resonance; (*c*) There are six points with $h \sim 0.5$ Å and $c \sim 3.5$ Å; five of these are Δ^2-carbapenems **13**–**17**[28,34–37] and one is a 1-oxa-Δ^2-penem **18**.[38] These molecules have unusually high *h* values for bicyclo[3.2.0] β-lactams not only because they are strained by fusion to a five-membered ring but also because amide resonance is inhibited by the enamine N-atom. In other words, these extremely strained molecules are more pyramidal than expected, but the *c* values are normal.

The gap (h = 0.25–0.35 Å and c = 3.3–3.8 Å) in the central portion of the *h*–*c* scatterplot (Fig. 6) has already been identified as a 'no-man's land' in Fig. 2 and corresponds to the structural gap between bicyclo[3.2.0] and bicyclo[4.2.0] β-lactams. Changes in molecular geometry between these groups

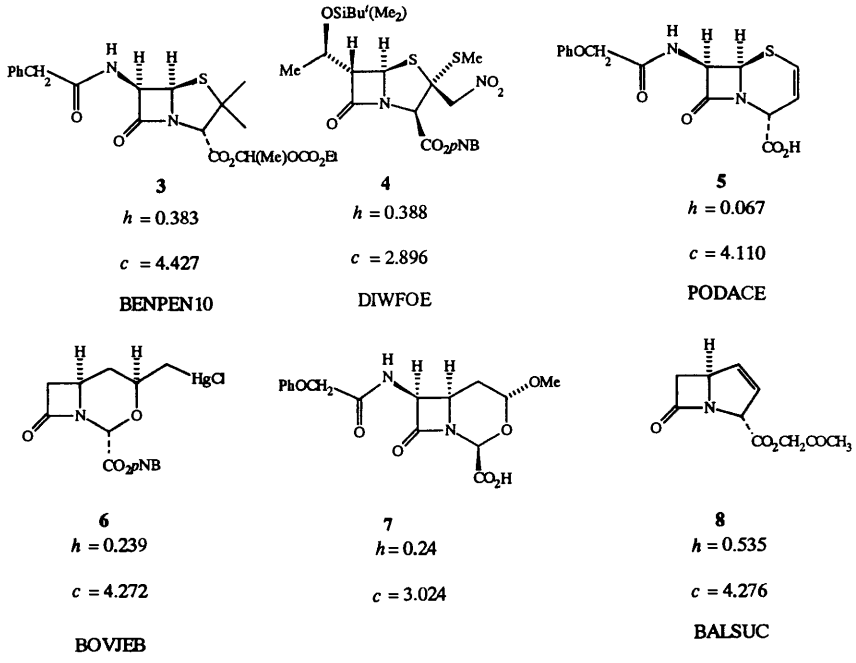


Fig. 8 Specific compounds in this study with REFCODE, *h* and *c* values. For **7**, the geometrical parameters have been derived in reference 14.

of molecules is so severe that a discrete jump in *h* values results, creating the empty central region in the scatterplot. To fill this region, compounds with intermediate pyramidality must be designed by appropriate functional group and/or heteroatom modification, although at present there are no such examples. It should be interesting to investigate the synthesis, 3D-structure and biological activity of such novel hybrid molecules, **31** and **32** (Fig. 10).

Of the remaining 34 structures (114 original less 80 active), three categories were defined according to the following skeletons: β-lactams fused to five-membered rings (bicyclo-[3.2.0] β-lactams, 10 hits); β-lactams fused to six-membered rings (bicyclo[4.2.0] β-lactams, 14 hits) and monocyclic β-lactams (6 hits). These 30 mono- and bi-cyclic β-lactams were classified as inactive because these skeletons are not known to be of significant clinical importance. The *h*–*c* scatterplot of these 30 inactive structures (Fig. 7) shows that less than half the points (13) lie along the 'active' diagonal. Of these 13 structures, one is a monocyclic β-lactam, seven are bicyclo[4.2.0] β-lactams and the rest are bicyclo[3.2.0] β-lactams.[28,39–50] The chemical structures of these compounds are given in Fig. 8 (compound **8**) and Fig. 9(*b*) (**19–30**). Compound **19** has a normal *c* distance (3.75 Å) and a somewhat high *h* value (0.21 Å) for a monocyclic β-lactam. This is because the bulky β,β-dimethylacrylic acid substituent on the nitrogen atom is moved out of the β-lactam ring plane so as to avoid steric crowding with the adjacent ethylsulfinyl residue. Compounds **20–26**, which lie along the active diagonal (Fig. 7) in the region h = 0.0–0.3 Å and c = 3.0–3.5 Å, are a varied lot of bicyclic β-lactams with no particular commonality in the nature of the four-atom tether or the substitution pattern on the rings. That these seven β-lactams have low *h* and low *c* values may be coincidental. In contrast, the cluster of five bicyclo[3.2.0] β-lactams **8**, **27–30** (h ~ 0.5 Å and c = 4.3–4.4 Å) which lie along the upper diagonal of Fig. 7 indeed have some common structural features, the most obvious being a highly folded bicyclic skeleton with an *exo*-oriented carboxy group. These features are desirable for activity but as was alluded to earlier, the lack of biological activity cannot (as yet) be rationalized solely with the joint *h* and *c* correlation for all compounds. The remaining points are scattered all over the off-diagonal region. To complete the book-keeping, it was observed that the four residual compounds with miscellaneous structures stand alone. Details of these compounds are given in the Supplementary Information.

That Figs. 6 and 7 are complementary in nature, with active compounds lying along the *h*–*c* diagonal and inactive ones in the off-diagonal region, confirms that a simultaneous examination of *h* and *c* parameters is crucial in establishing structure–activity relationships in β-lactam antibiotics. Examination of Figs. 6 and 7 also reveals why a consideration of either of these parameters in isolation often provides only a poor idea of biological activity. In summary, the correlation between these two independent structural parameters together with biological activity is unambiguous and hence the degree of reliability in its use should be far superior.

Model for binding of β-lactams to PBPs

It is known from X-ray crystal structures of carboxypeptidases, transpeptidases and β-lactamases that penicillin-recognizing enzymes contain the conserved sequence Ser–X–X–Lys in their active site.[5,51] The importance of non-covalent interactions in facilitating enzyme-substrate recognition is also crystallographically well-documented.[5–8] In general, it is accepted that there is a three-point binding of the enzyme receptor and the β-lactam molecule. Recognition occurs *via:* (*a*) the incipient attack by the serine –OH group on the carbonyl group of the lactam; (*b*) electrostatic interaction between the lysine NH_3^+ of the enzyme protein and the CO_2^- group at C(3); (*c*) hydrogen bonding by the C(6) amide N–H group with a valine carbonyl group (penicillin numbering). For effective binding, these specific interactions must also be accompanied by a good van der Waals fit of the drug molecule in the receptor pocket.

It is now instructive to examine Fig. 11 which is a stereo diagram showing the superposition of the β-lactam–C–CO_2 fragment in 12 active compounds chosen from Fig. 6. Compounds close to the 'active' diagonal were chosen but other than this, the choice was largely random and Fig. 11 therefore contains both high *h* and high *c* and low *h* and low *c* compounds. Fig. 11 shows that the lactam carbonyl groups and the carboxylate groups are bunched in their specific regions. So, this depiction is a good pharmacophore model[52] for two-point binding of the β-lactam with the enzyme. It is significant to note that even with the amide carbonyl and carboxyl groups thus tethered to the enzyme, there is not much variation in

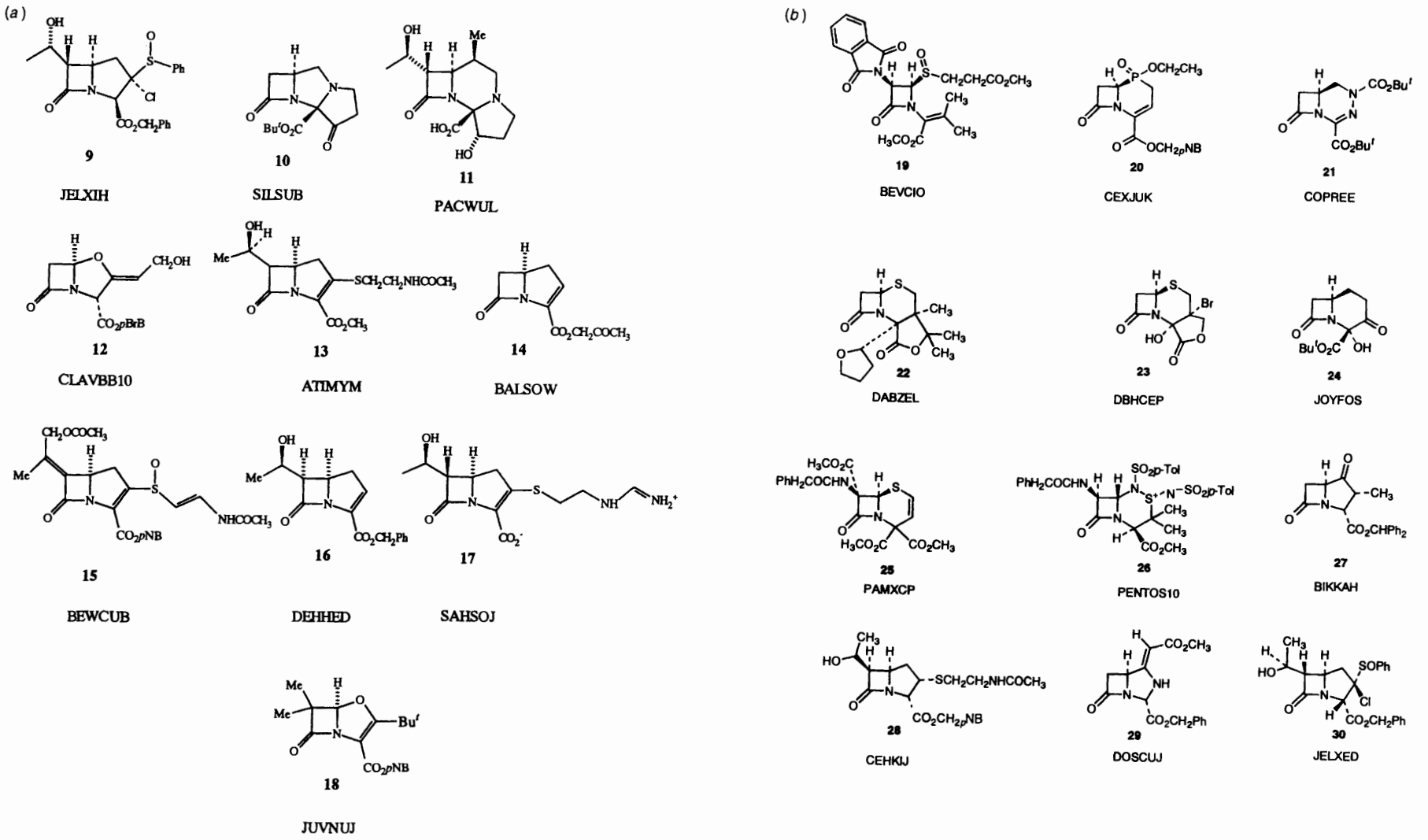
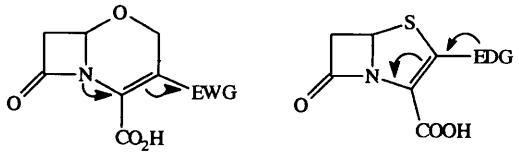


Fig. 9 (*a*) Structures of some outlier compounds in Fig. 6 with REFCODE. (*b*) Structures of compounds lying on the 'active' diagonal in Fig. 7.



31 *h* increases **32** *h* decreases

Fig. 10 Hypothetical molecules which may lie in the 'no-man's-land' in the *h*–*c* scatterplots (medium *h* and medium *c*)

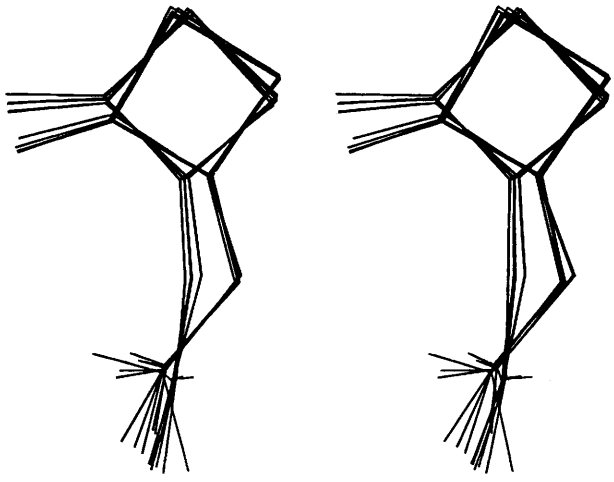


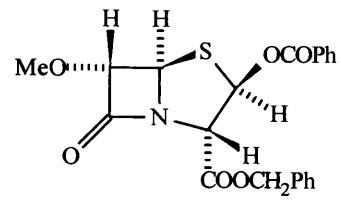
Fig. 11 Superposition stereoplot of 12 active structures chosen from Fig. 6. The compounds were chosen as near as possible to the diagonal.

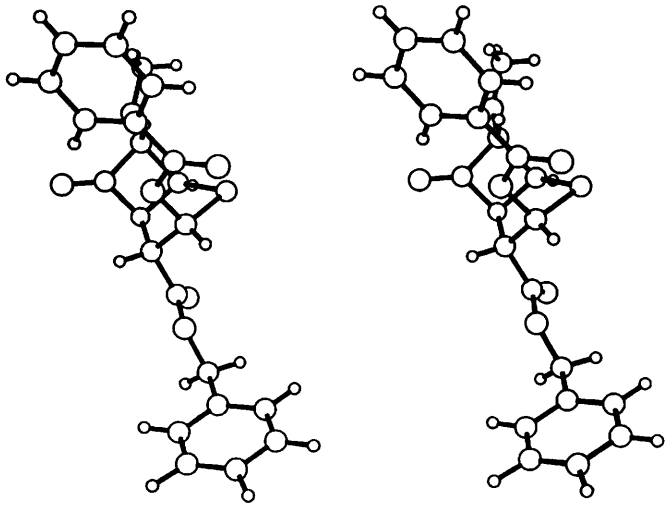
the positioning of the four-membered lactam ring and the connecting –C– linkage in these 12 compounds. Both high *h* and high *c* and low *h* and low *c* structures are tightly grouped spatially and it is not difficult to visualise that even with the fused 5- or 6-membered ring and attendant substitution filled in, no great spatial variation is expected. In contrast, when an extra 'wrong' structure was intentionally added in Fig. 12, its outlier status is easily identified. Though penam **21** (FAHRAH) is a high *h* and high *c* compound, Fig. 12(*a*) shows that the carboxylate group is in the 'inactive' pseudo-axial conformation[14] and Fig. 12(*b*) shows that it cannot be superposed easily with the active compounds.

Fig. 13 is the corresponding superposition plot of 12 'off-diagonal', inactive compounds selected from Fig. 7. The lactam carbonyl and the carboxylate groups have again been bunched as close as possible (the latter groups perhaps not so effectively as in Fig. 11), because this is the prerequisite for binding and, eventually, biological activity. However, in the attempt at such a superpositioning, a considerable variation in the positioning of the lactam ring and the –C– linkage is observed. The average lactam ring planes for the low *h* and high *c* compounds and the high *h* and low *c* compounds are sharply inclined. Further, the positions of the –C– linkage now occur in distinct and widely-distributed spatial regions and it is clear that when the entire molecular structure of the antibiotic is filled in, a considerable spatial variation will result.

Noting that spatial factors are important here and that torsion angles are better single-parameter descriptors of shape than distances, we examined torsional parmeters about non-bonded vectors in an attempt to more precisely quantify this shape parameter which convolutes the information supplied by the Woodward and Cohen parameters. Such an analysis is relevant because it needs quite a number of distances to equivalence a single torsional parameter, thus increasing, often needlessly, the number of parameters to be taken into account

(*a*)





(*b*)

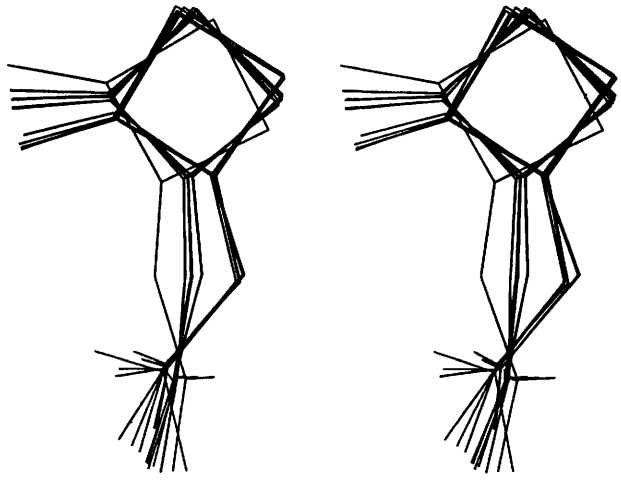


Fig. 12 (*a*) Structure and stereoview of molecule **21** (refcode FAHRAH). Though nominally with an active skeleton, **21** has the 'inactive' pseudo-axial conformation. (*b*) Plot in Fig. 11 with penem **21** additionally included.

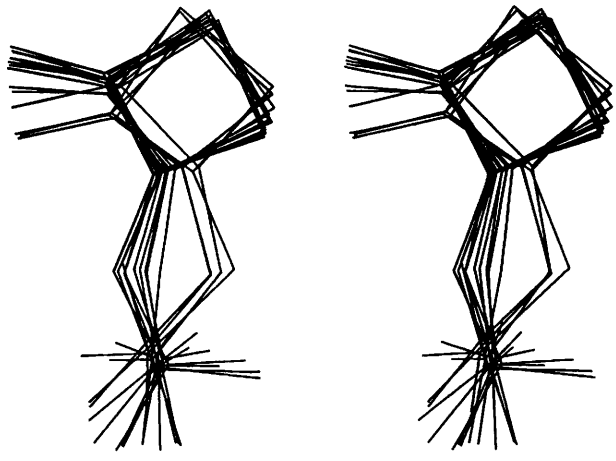


Fig. 13 Superposition stereoplot of 12 inactive structures chosen from Fig. 7. The compounds were chosen randomly.

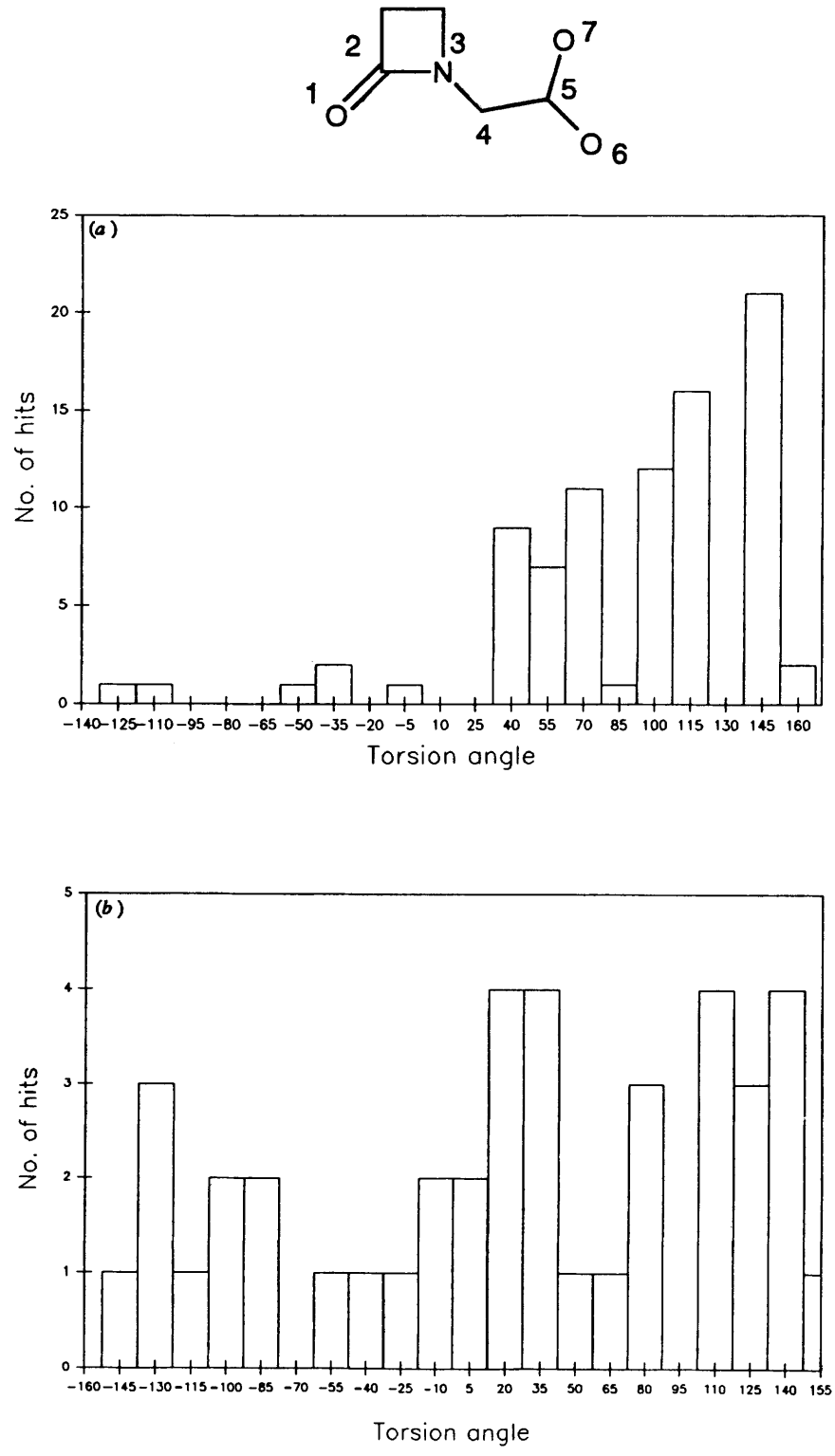


Fig. 14 (*a*) Histogram of the 1–3–4–5 torsional angle values for the 80 active β-lactams. (*b*) Histogram of the 1–3–4–5 torsional angle values for the remaining 34 structures.

in multivariate structure correlation analysis. A number of torsional angles were examined. Fig. 14 is a histogram of the torsional angle 1–3–4–5 for active and inactive compounds; note that the shape of the β-lactam is estimated by using the non-bonded vector 1–3. The figure shows that this torsional angle lies in the region 30–160° for the active compounds while the angles for the inactive compounds are scattered in the entire region −160 to +155°. Further studies on this aspect are in progress.

Our observations concerning Figs. 11, 12, 13 and 14 underscore the fact, recently elaborated by Hahn,[52] that pharmacophore models tend to be geometrically underconstrained, because molecules that fit the model can still be inactive because of additional regions of the molecule that are located in sterically unfavourable locations. In our context, all β-lactams can be spatially positioned so that the lactam carbonyl group and the carboxylate group can align, to a greater or lesser extent, with relevant complementary regions of

the enzyme, but an overall fit of the molecule in the receptor cavity is only possible for the high h and high c and low h and low c compounds.

Conclusions

We have studied a large number of β-lactam antibiotic crystal structures retrieved from the CSD to investigate the range of favourable values of the empirical structural parameters defined by Woodward (h) and Cohen (c) for antibacterial activity.[11,14] We have found that the range of h values in clinically useful antibiotics is in agreement with data collected manually over the last 15 years. However, the range of c distances for active geometries is updated to 3.0–4.5 Å from the earlier limit of 3.0–3.9 Å.[14] Additionally, we have found that a joint analysis of h and c parameters provides a better, though still empirical, correlation of structure with biological activity. Since the two parameters are independent and their combined consideration leads to better predictions of activity, their convolution defines a third and more significant parameter. This is shown to be an overall shape factor, which is quantified as a torsional angle around a non-bonded vector. Active molecules with high h and high c or low h and low c values are able to adopt a nearly similar conformation in the β-lactam-C-carboxylate region while the inactive molecules with high h and low c and low h and high c show much spatial variation. These results indicate that the receptor cavity in penicillin-binding proteins has a well-defined geometry and that shape recognition, without much induced fit,[53] is an important prerequisite for binding of β-lactam antibiotics and subsequent biological activity. The importance of overall shape complementarity between the β-lactam drug and its receptor protein as revealed by the active diagonal on the h–c scatterplot is expected to be a useful guide in the search of newer fourth-generation antibiotics.

Experimental

Cambridge Structural Database (CSD) analysis

Data were retrieved from the 1994 update (Version 5.05) of the CSD (109 816 entries) for all the ordered crystal structures with an exact match between chemical and crystallographical connectivity and containing at least one occurrence of the β-lactam fragment **2**. Structures with R-values greater than 0.100 were rejected. Duplicate hits (identified by the same REFCODE) were removed manually by eliminating all but the structure with the lowest R-value in each case. A total of 114 structures was retrieved. Geometrical calculations were performed on the retrieved data to calculate h, c, r and ΣN using QUEST3D-GSTAT, an automatic graphics search program. These values are provided in the Supplementary Information. The superposition routines [Figs. 11, 12(*b*) and 13] were also taken from the CSD package. Calculations were performed on a Microvax 3300 computer.

Acknowledgements

We thank C. Asha Jyothi and R. Kulsum for their assistance in the literature survey. One of us (K. B.) thanks the UGC for fellowship support.

References

1 *Chemistry and Biology of β-Lactam Antibiotics*, eds. R. B. Morrin and M. Gorman, Academic Press, New York, 1982, vols. 1–3.
2 *The Organic Chemistry of β-Lactams*, ed. G. I. Georg, VCH, New York, 1993.
3 W. Dürckheimer, J. Blumbach, R. Lattrell and K. H. Scheunemann, *Angew. Chem. Int. Ed. Engl.*, 1985, **24**, 180.
4 A. G. Brown, *Pure Appl. Chem.*, 1987, **59**, 475.
5 J. A. Kelly, J. R. Knox, P. C. Moews, G. J. Hite, J. B. Bartolone, H. Zhao, B. Joris, J.-M. Frere and J.-M. Ghuysen, *J. Biol. Chem.*, 1985, **260**, 6449.
6 J. A. Kelly, J. R. Knox, H. Zhao, J.-M. Frere and J.-M. Ghuysen, *J. Mol. Biol.*, 1989, **209**, 281.
7 J. R. Knox and P. C. Moews, *J. Mol. Biol.*, 1991, **220**, 435.
8 C. Oefner, A. D'Arcy, J. J. Daly, K. Gubernator, R. L. Charnas, I. Heinze, C. Hubschwerlen and F. K. Winkler, *Nature*, 1990, **343**, 284.
9 N. H. Georgopapadakou, *Antimicrob. Agents Chemother.*, 1993, **37**, 2045.
10 B. G. Spratt, *Science*, 1994, **264**, 388.
11 R. B. Woodward, *Phil. Trans. R. Soc. Lond.*, 1980, **B289**, 239 and references cited therein.
12 D. B. Boyd, in *Chemistry and Biology of β-Lactam Antibiotics*, eds. R. B. Morin and M. Gorman, Academic Press, New York, 1982, vol. 1, p. 437.
13 R. M. Sweet and L. F. Dahl, *J. Am. Chem. Soc.*, 1970, **92**, 5489.
14 N. C. Cohen, *J. Med. Chem.*, 1983, **26**, 259.
15 D. D. Keith, J. Tengi, P. Rossman, L. Todardo and M. Weigele, *Tetrahedron*, 1983, **39**, 2445.
16 F. H. Allen, J. E. Davies, J. J. Galloy, O. Johnson, O. Kennard, C. F. Macrae, E. M. Mitchell, G. F. Mitchell, J. M. Smith and D. G. Watson, *J. Chem. Inf. Comput. Sci.*, 1991, **31**, 187.
17 G. R. Desiraju, *Crystal Engineering. The Design of Organic Solids*, Elsevier, Amsterdam, 1989, p. 63.
18 G. Klebe, in *Structure Correlation*, eds. H.-B. Bürgi and J. D. Dunitz, VCH, Weinheim, 1994, vol. 2, p. 544.
19 W. L. Duax, J. F. Griffin and D. Ghosh, in *Structure Correlation*, eds. H.-B. Bürgi and J. D. Dunitz, VCH, Weinheim, 1994, vol. 2, p. 605.
20 O. A. Mascaretti, O. A. Roveri and G. O. Danelon, in *Recent Progress in the Chemical Synthesis of Antibiotics and Related Microbial Products*, ed. G. Lukacs, Springer, Berlin, 1993, vol. 2, p. 677.
21 R. Taylor and F. H. Allen, in *Structure Correlation*, eds. H.-B. Bürgi and J. D. Dunitz, VCH, Weinheim, 1994, vol. 1, p. 111.
22 R. Busson, P. J. Claes and H. Vanderhaeghe, *J. Org. Chem.*, 1976, **41**, 2556.
23 I. Csöregh and T.-B. Palm, *Acta. Crystallogr., Sect. B*, 1977, **33**, 2169.
24 A. Bedeschi, F. Bélanger-Gariepy, S. Hanessian and F. Brisse, *Acta. Crystallogr., Sect. C*, 1986, **42**, 442.
25 M. L. Phillips, R. Bonjouklian, N. D. Jones, A. H. Hunt and T. K. Elzey, *Tetrahedron Lett.*, 1983, **24**, 335.
26 J. A. Gleason, T. F. Buckley, K. G. Holden, D. B. Bryan and P. Siler, *J. Am. Chem. Soc.*, 1979, **101**, 4730.
27 N. C. Cohen, P. Colin and G. Lemoine, *Tetrahedron*, 1981, **37**, 1711.
28 H. R. Pfaendler, J. Gosteli, R. B. Woodward and G. Rihs, *J. Am. Chem. Soc.*, 1981, **103**, 4526.
29 C. L. Branch and M. J. Pearson, *J. Chem. Soc., Perkin Trans. 1*, 1986, 1077.
30 J. H. Bateson, P. M. Roberts, T. C. Smale and R. Southgate, *J. Chem. Soc., Perkin Trans. 1*, 1990, 1541.
31 H. H. Wasserman, S. L. Henke, E. Nakanishi and G. Schulte, *J. Org. Chem.*, 1992, **57**, 2641.
32 H. H. Wasserman, S. L. Henke, P. Luce, E. Nakanishi and G. Schulte, *J. Org. Chem.*, 1990, **55**, 5821.
33 A. G. Brown, D. F. Corbett, J. Goodacre, J. B. Harbridge, T. T. Howarth, R. J. Ponsford, I. Stirling and T. J. King, *J. Chem. Soc., Perkin Trans. 1*, 1984, 635.
34 G. Albers-Schönberg, B. H. Arison, O. D. Hensens, J. Hirshfield, K. Hoogsteen, E. A. Kaczka, R. E. Rhodes, J. S. Kahan, F. M. Kahan, R. W. Ratcliffe, E. Walton, L. J. Ruswinkle, R. B. Morin and B. G. Christensen, *J. Am. Chem. Soc.*, 1978, **100**, 6491.
35 N. Tsuji, K. Nagashima, M. Kobayashi, J. Shoji, T. Kato, Y. Terui, H. Nakai and M. Shiro, *J. Antibiot.*, 1982, **35**, 24.
36 J. H. Bateson, A. M. Quinn, T. C. Smale and R. Southgate, *J. Chem. Soc., Perkin Trans. 1*, 1985, 2219.
37 R. W. Ratcliffe, K. J. Wildonger, L. di Michele, A. W. Douglas, R. Hajdu, R. T. Goegelman, J. P. Springer and J. Hirshfield, *J. Org. Chem.*, 1989, **54**, 653.
38 H. R. Pfaendler, W. Hendel and U. Nagel, *Z. Naturforsch., Teil B*, 1992, **47**, 1037.
39 M. D. Bachi, A. Gross and F. Frolow, *J. Org. Chem.*, 1982, **47**, 765.
40 H. Satoh and T. Tsuji, *Tetrahedron Lett.*, 1984, **25**, 1733.
41 C. M. Pant, R. J. Stoodley, A. Whiting and D. J. Williams, *J. Chem. Soc., Chem. Comm.*, 1984, 1289.
42 N. K. Capps, G. M. Davies, P. B. Hitchcock and D. W. Young, *J. Chem. Soc., Chem. Comm.*, 1985, 843.

43 K. Neupert-Laves and M. Dobler, *Cryst. Struct. Commun.*, 1977, **6**, 153.
44 C. M. Gasparski, A. Ghosh and M. J. Miller, *J. Org. Chem.*, 1992, **57**, 3546.
45 E. F. Paulus, *Acta Crystallogr., Sect. B*, 1977, **33**, 108.
46 A. F. Cameron, I. R. Cameron, M. M. Campbell and G. Johnson, *Acta Crystallogr., Sect. B*, 1976, **32**, 1377.
47 R. L. Rosati, L. V. Kapili, P. Morrissey, J. Bordner and E. Subramanian, *J. Am. Chem. Soc.*, 1982, **104**, 4262.
48 T. Haneishi, M. Nakajima, N. Serizawa, M. Inukai, Y. Takiguchi, M. Arai, S. Satoh, H. Kuwano and C. Tamura, *J. Antibiot.*, 1983, **36**, 1581.
49 C. L. Branch and M. J. Pearson, *J. Chem. Soc., Perkin Trans. 1*, 1986, 1077.
50 J. H. Bateson, P. M. Roberts, T. C. Smale and R. Southgate, *J. Chem. Soc., Perkin Trans. 1*, 1990, 1541.
51 S. Wolff, *Can. J. Chem.*, 1994, **72**, 1014.
52 M. Hahn, *J. Med. Chem.*, 1995, **38**, 2080.
53 D. E. Koshland, *Angew. Chem. Int. Ed. Engl.*, 1994, **33**, 2375.

Paper 5/05796E
Received 1*st September* 1995
Accepted 1*st December* 1995