Journal of Computer-Aided Molecular Design, 8 (1994) 583–606
ESCOM

J-CAMD 254

A fast and efficient method to generate biologically relevant conformations

Gerhard Klebe* and Thomas Mietzner

BASF AG, Main Laboratory, Carl-Bosch Strasse, D-67056 Ludwigshafen, Germany

Received 5 January 1994
Accepted 25 March 1994

Key words: Conformational analysis of drug molecules; Biologically relevant conformations; Conformational preferences from crystal structures; Computer program; Comparison with protein-bound conformation

SUMMARY

Mutual binding between a ligand of low molecular weight and its macromolecular receptor demands structural complementarity of both species at the recognition site. To predict binding properties of new molecules before synthesis, information about possible conformations of drug molecules at the active site is required, especially if the 3D structure of the receptor is not known. The statistical analysis of small-molecule crystal data allows one to elucidate conformational preferences of molecular fragments and accordingly to compile libraries of putative ligand conformations. A comparison of geometries adopted by corresponding fragments in ligands bound to proteins shows similar distributions in conformation space. We have developed an automatic procedure that generates different conformers of a given ligand. The entire molecule is decomposed into its individual ring and open-chain torsional fragments, each used in a variety of favorable conformations. The latter ones are produced according to the library information about conformational preferences. During this building process, an extensive energy ranking is applied. Conformers ranked as energetically favorable are subjected to an optimization in torsion angle space. During minimization, unfavorable van der Waals interactions are removed while keeping the open-chain torsion angles as close as possible to the experimentally most frequently observed values. In order to assess how well the generated conformers map conformation space, a comparison with experimental data has been performed. This comparison gives some confidence in the efficiency and completeness of this approach. For some ligands that had been structurally characterized by protein crystallography, the program was used to generate sets of some 10 to 100 conformers. Among these, geometries are found that fall convincingly close to the conformations actually adopted by these ligands at the binding site.

INTRODUCTION

Flexible molecules often exhibit several conformations of nearly equal energy. Which one of these is adopted under any given conditions may depend on the particular interactions with the surrounding environment and its relative energy content. The systematic sampling of conforma-

*To whom correspondence should be addressed.

tions is often intractable, because of the complexity of the problem. Even for a simple case like *n*-hexane, with three C-C-C-C torsion angles (omitting rotations about the bonds to the terminal methyl groups), a systematic search in 30° intervals will result in $12^3 = 1728$ arrangements (or 864 if symmetry is regarded). These geometries would have to be further examined, for example by molecular mechanics, to elucidate into how many local minima these starting structures will converge. For the prediction of binding properties only those conformations that can possibly occur at the binding site of a protein are relevant. Therefore, methods are required that can generate those conformations in an efficient and reliable manner.

In general, the conformational properties of a molecule are determined by its intramolecular interactions. In a series of studies [1] it has been shown that, depending on the actual substitution pattern, the accessible conformations of local torsional fragments (such as n-membered rings, side chains at aromatic moieties, or open-chain portions) cluster about the local minima of the idealized unsubstituted reference fragments. For example, in a sterically unperturbed situation, a connection between two sp^3-hybridized atoms shows local minima at approximately $\pm 60°$ and 180°; those joining an sp^2- and sp^3-hybridized center possess potential minima at every 30°. In addition to the intramolecular forces, conformations can be perturbed substantially by intermolecular interactions with the surrounding environment. The question can be raised whether both the intramolecular force field and the interactions with the environment restrain the conformational multiplicity of a given molecule. If so, it can be expected that the local torsional fragments occur only with a limited number of preferred angles. Using such information on preferred angles in a knowledge-based approach should allow one to map the relevant parts of conformation space using much less starting structures compared to, for example, a systematic variation of each rotatable bond in small incremental steps.

In order to obtain some evidence for the question put forward on an experimental basis, the geometries of ligands observed in crystal structures of ligand–protein complexes were compared to those obtained from small-molecule crystal structure analyses and computational methods [2]. Crystallographic data are of particular relevance for the present problem, since conformations of molecules adopted in a crystal correspond to geometries exhibited in an anisotropic molecular environment. These conditions should allow one to probe simultaneously the influence and importance of both the intramolecular force field and the typical distortions arising from interactions between functional groups of neighboring molecules to form a supramolecular assembly.

In summary, the comparative study [2] of protein and small-molecule data revealed the following conclusions: reliance on just the crystal structures of a ligand itself may not, in most cases, be an infallible and reliable indicator of a biologically active conformation. A similar observation has recently been reported by Ricketts et al. [3]. Perhaps, this approach may be justified to some extent for rather nonpolar and rigid ligands, where the environment is composed of many isotropically distributed atom/atom potential contributions. In such cases, it can be assumed that the adopted conformation is mainly determined by intramolecular forces and that no significant perturbations arise from strong directional interactions with the crystal or active-site environment. Presumably, this assumption is also valid for hydrophobic portions of a ligand skeleton. For polar ligands, embedded into a network of strongly directional hydrogen bonds, the environment clearly influences the adopted conformation. In these cases, the analysis was focused on the study of torsional fragments in the molecules. It was found that the statistical evaluation of conformational preferences of torsional fragments, observed in small-molecule crystal structures, provides

important information about conformations relevant at a receptor site [2]. Conformational libraries are obtained that possess some predictive ability to generate relevant binding geometries. Compared to a systematic scanning of the rotatable bonds in, e.g., 30°-steps, they give rise to much less conformational multiplicity. Thus, a substantial speed-up of the conformational search can be expected.

However, the analysis also showed that the most frequently occurring geometry in small-molecule crystal structures is not necessarily the one adopted in a protein environment. The less populated conformations have to be considered as well. At the protein-binding site, geometries can be adopted which deviate substantially from a local energy minimum geometry of the isolated molecules in order to optimize their interactions with the neighboring active-site residues. This holds especially for substrates, where an enzyme induces specific distortions in order to prepare the substrate molecule for the transition state of a chemical reaction.

The present paper describes a knowledge-based approach toward conformational analysis. This 'knowledge', especially on open-chain portions, is not stored in terms of 'rules', but incorporated into an extended set of conformational libraries which are directly derived from crystal data by a statistical evaluation of conformational preferences. They are used to generate conformations of small-molecule ligands and to avoid unfavorable close contacts. The present approach is entirely based on experimental data and empirical concepts. Thus, its relevance can only be demonstrated by applications that produce results similar to experimental evidence. Accordingly, to assess how well the generated conformations are distributed in conformation space, a comparison with data obtained from experimental sources is performed. For some typical protein ligands we check whether one of the generated conformations is structurally close to the geometry actually adopted at the binding site of the protein.

PROGRAM FOR AUTOMATIC CONFORMATIONAL ANALYSIS USING LIBRARIES OF TORSION ANGLES OBTAINED FROM CRYSTAL DATA

The information on conformational preferences has been incorporated into the automatic procedure MIMUMBA for the generation of conformations relevant for the structural requirements at the binding site of a protein. A drug molecule under consideration can be analyzed in terms of its conformational properties. The molecule has to be supplied in any arbitrary conformation with standard bond lengths and bond angles, e.g. model-built from a crystal structure or generated by a 2D/3D conversion program such as CONCORD [4] or CORINA [5]. By assigning rotatable bonds, the molecule is split into rigid fragments or flexible ring portions. This partitioning is performed either automatically by the program or according to assignments given by the user. To enhance a fast search for fairly large and extended molecules, optionally some selection rules can be applied. They consider less conformational flexibility for molecular portions at the termini (since they influence less the general shape of a molecule) and exploit an energy ranking (see below) attributed to the various conformations of the individual fragments.

Analysis of the ring portions

The program treats ring and open-chain portions separately. A detailed analysis of the flexible ring portions is performed according to the program SCA [6], which has been adapted and interfaced to MIMUMBA. The cyclic fragments which may be constructed from several fused

rings are formally separated into the smallest possible ring units (up to seven-membered rings). For these elementary units, low-energy conformations are taken from a table with predefined ring templates derived from experimental data (e.g., crystal structures). A relative energy content is assigned to the various templates [6]. Any conformational multiplicity due to pseudorotation or inversion of each elementary ring is considered.

Subsequently, the program recomposes the individual elementary ring units in the various conformations into the entire polycyclic ring system, with regard to the required stereochemistry. An energy ranking is performed according to the relative energy content of the individual building blocks and additional contributions are considered for the relative positioning of substituents at the rings. For the fusion of elementary rings, a set of rules on preferable torsion angles has been defined. The angles achieved during this fusion process are energetically ranked. The further they deviate from a predefined optimal value, the less favorable the obtained ring system is ranked. Steric overlap between atoms in 1,4-, 1,5- and 1,6-positions is checked and occasionally an energy 'penalty' is assigned [6]. Finally, the program rejects conformations of the composed rings if the assigned energy ranking rises beyond a predefined limit.

Analysis of the open-chain portions

Dihedral angles for rotatable bonds in the open-chain portions were selected from a database which contains the results of statistical analyses on the conformational preferences of torsional fragments. Appropriate data have been retrieved from the Cambridge Structural Database (CSD) [7]. Different versions of the CSD between 1987 and 1992 have been used for data retrieval. Structures containing the torsional fragment of interest were matched by a connectivity search using the QUEST program in the CSD program suite [7]. Depending on the number of hits obtained for a particular query, those structures flagged in the database as being erroneous or distorted were rejected from the analysis and an R-factor rejection criterion of $R_F < 0.1$ was applied. Subsequent geometry evaluations have been performed with the GSTAT program [7] by means of univariate statistics. The connectivity of the probe fragment was coded according to the QUEST run. However, since GSTAT solely operates on crystallographic atomic coordinates, the chemical connectivity present in QUEST has to be generated from the distance information. Some misclassification of torsional fragments could occur if not all of the hydrogen atoms had been located in the diffraction study. To reduce possible misclassifications, if required, distance tests were applied using standard values obtained from crystal data [8].

Considering the local symmetry of the fragments, all torsion-angle values have been projected into the same asymmetric unit of the conformation space [9]. In principle, the atoms forming a torsional fragment can give rise to molecular asymmetry. During data retrieval from the CSD, no control on the absolute configuration was performed. As a consequence, the univariate statistics were evaluated as if obtained from a racemic data sample. Figure 1 gives, as an example, the distribution of torsion angles for an anisole-type fragment (left). The retrieved fragments were classified according to their substitution pattern on the ortho-positions with respect to the OCH_3 group: ortho-unsubstituted (both H), mono ortho-substituted (one ortho-H, one ortho non-hydrogen) and ortho-disubstituted derivatives (both ortho non-hydrogens). For symmetrically substituted examples, the asymmetric unit corresponds to the range 0–90° and the distribution repeats mirror-symmetrically after 90°. The periodicity of the total distribution is 180°. For the mono-substituted derivatives, a mirror-symmetric distribution is achieved after 180°. Similarly,

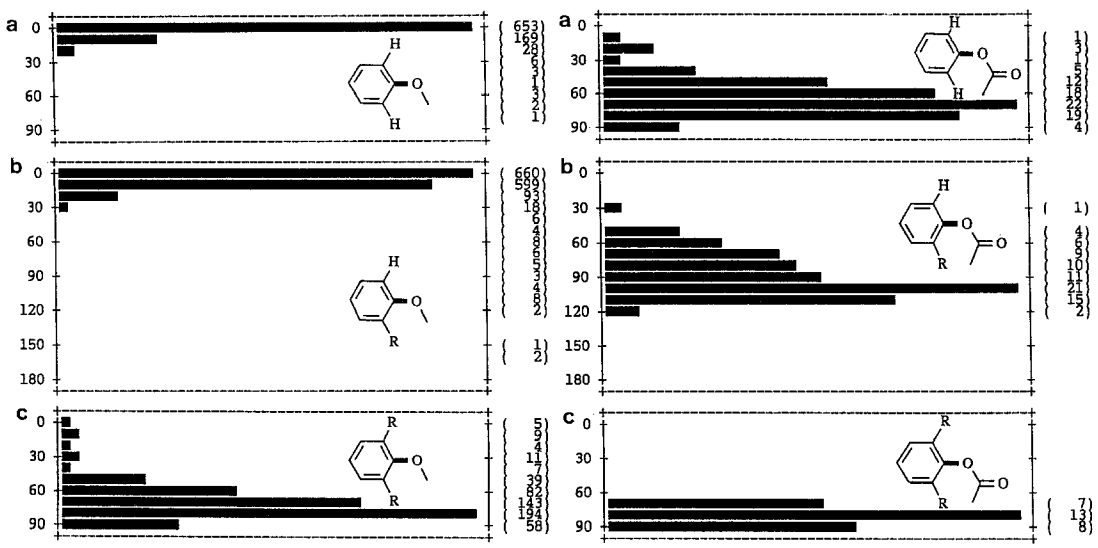


Fig. 1. Univariate statistics of torsion angles in different fragments, as observed in small-molecule crystal structures. Left: anisole-type; right: phenoxyester fragments; with (a) both ortho-positions occupied by H; (b) one ortho = H, one ≠ H; (c) both ortho non-hydrogen.

each torsional fragment included into the database has been analyzed and subsequently classified according to its symmetry and periodicity. In the database, each fragment is characterized by a sequence of atom and bond types (similar to the SYBYL-type notation [10], e.g., Car ar Car 1 O3 1 C3 for the anisole example). Any special substitution patterns, e.g. non-hydrogen substituents at the ortho-positions of aromatic rings or at a double bond, or special substituents (e.g. O2) at the terminal end of a fragment have been coded together with the appropriate symmetry and periodicity. Figures 1 and 2 show the univariate statistics of some additional torsional fragments: phenoxyesters, open-chain alkanes, amines and ethers. At present, the database comprises 216 torsional fragments. We believe that this collection covers the most important torsion angles occurring in typical organic molecules. The examples given in Figs. 1 and 2 are coded in the database by the sequence listed in Fig. 3. In addition, their symmetry and periodicity are given together with an assigned energy ranking (see below). The results of the different univariate statistics have been deposited with the Fachinformationszentrum Karlsruhe*. On request, they can be ordered using the deposition code CSD-58048 [11].

Energy ranking of the generated conformers

The various histograms stored for the different torsional fragments have been ranked energetically in steps of 30° on a relative energy scale (between 0 and 10 kJ). This ranking is based on literature data, semiempirical model calculations and on the relative frequency of the various experimentally observed conformations (see Fig. 3). After classifying the open-chain fragments of the probe molecule, the program retrieves matching fragments from the database. If more than one entry in the database could be found for a particular torsion axis, the intersection of the

*Fachinformationszentrum Karlsruhe, D-76344 Eggenstein-Leopoldshafen, Germany.

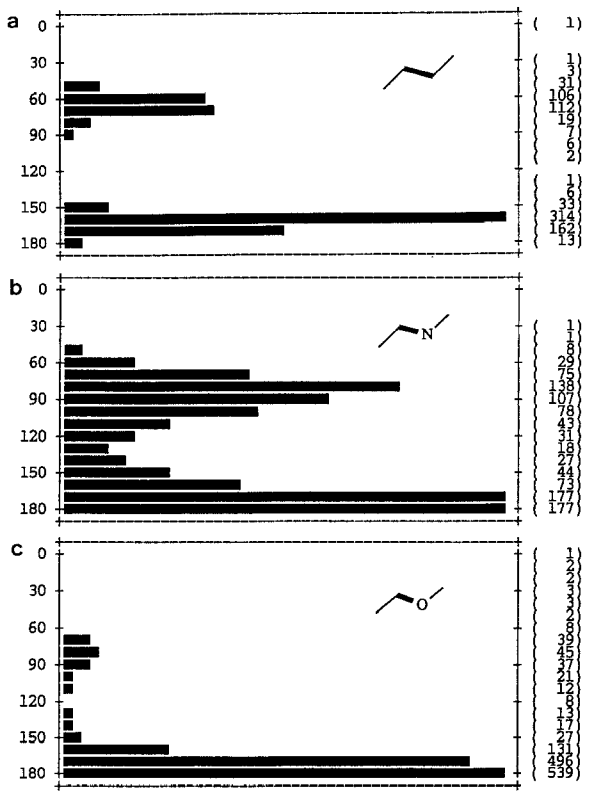


Fig. 2. Univariate statistics of torsion angles in different fragments, as observed in small-molecule crystal structures. (a) Open-chain alkanes; (b) amines; and (c) ethers.

corresponding angles is used. The energy ranking of the remaining angles is assigned to the maximal values occurring in one of the matched probe fragments. Similar to the study of cyclic ring portions, the program composes the individual fragments into the total molecule, using preferably those open-chain torsion angles which are ranked favorable.

While merging all ring and open-chain fragments in their various conformations, the individual energy contributions are added and again a 'penalty' is assigned if fragments overlap. This penalty function is set proportional to the overlapping volumes between the various fragments being combined. It is calculated by the analytical function expressing the common volume of mutually penetrating spheres. An overlap of 1 $Å^3$ is assigned to a penalty of 1 kJ. The ranking of every possible combination of fragments in their various conformations is determined in a breadth-first manner. At first, all substructures composed by two adjacent fragments are examined. Then, three basic fragments are combined and, subsequently, the substructures are gradually increased. Suppose that we want to compute the ranking of a substructure a-B-c, where B is the central moiety in a substructure of linear sequence and a and c are fragments connected to either one of the terminal ends of B. The total ranking of this substructure can be calculated from the following pairwise contributions:

$$E_{a\text{-}B\text{-}c} = E_{a\text{-}B} + E_{B\text{-}c} - E_B + E_{ac}$$

```
entry 1: ortho unsubstituted anisoles (s. Fig. 1, left)
CAR(H) AR CAR(H) 1  O3 1  C3
      0.0   30.0   60.0   90.0                          periode:  180.0
        0      3      5      7                          symmetry:   90.0
   0  10  20  30  40  50  60  70  80  90
 653 169  28   6   3   1   3   2   1   0
-------------------------------------------------------------------------
entry 2: ortho monosubstituted anisoles (s. Fig. 1, left)
CAR(H) AR CAR(R) 1  O3 1  C3
      0.0   30.0   60.0   90.0  120.0  150.0           periode:  360.0
        0      1      3      3      6      9           symmetry:  180.0
  0  10  20  30  40  50  60  70  80  90 100 110 120 130 140 150 160 170 180
660 599  93  18   6   4   8   6   5   3   4   8   2   0   0   1   2   0   0
-------------------------------------------------------------------------
entry 3: ortho disubstituted anisoles (s. Fig. 1, left)
CAR(R) AR CAR(R) 1  O3 1  C3
       0.0   30.0   60.0   90.0                        periode:  180.0
         8      6      3      0                        symmetry:   90.0
   0  10  20  30  40  50  60  70  80  90
   5   9   4  11   7  39  82 143 194  58
-------------------------------------------------------------------------
entry 4: ortho unsubstituted phenoxyesters (s. Fig. 1, right)
CAR(H) AR CAR(H) 1  O3 1  C2(2O2)
     60.0                                              periode:  180.0
        0                                              symmetry:   90.0
   0  10  20  30  40  50  60  70  80  90
   0   1   3   1   5  12  18  22  19   4
-------------------------------------------------------------------------
entry 5: ortho monosubstituted phenoxyesters (s. Fig. 1, right)
CAR(R) AR CAR(H) 1  O3 1  C2(2O2)
     90.0                                              periode:  360.0
        0                                              symmetry:  180.0
  0  10  20  30  40  50  60  70  80  90 100 110 120 130 140 150 160 170 180
  0   0   0   1   0   4   6   9  10  11  21  15   2   0   0   0   0   0   0
-------------------------------------------------------------------------
entry 6: ortho disubstituted phenoxyesters (s. Fig. 1, right)
CAR(R) AR CAR(R) 1  O3 1  C2(2O2)
     90.0                                              periode:  180.0
        0                                              symmetry:   90.0
   0  10  20  30  40  50  60  70  80  90
   0   0   0   0   0   0   0   7  13   8
-------------------------------------------------------------------------
entry 7: open-chain alkanes (s. Fig. 2, upper row)
CANY 1  C3(H)(H) 1  C3(H)(H) 1  CANY
     60.0  180.0                                       periode:  360.0
        4      0                                       symmetry:  180.0
  0  10  20  30  40  50  60  70  80  90 100 110 120 130 140 150 160 170 180
  1   0   0   1   3  31 106 112  19   7   6   2   0   1   6  33 314 162  13
-------------------------------------------------------------------------
entry 8: open-chain amines (s. Fig. 2, central row)
CANY 1  NANY(H) 1  C3(H)(H) 1  CANY
     90.0  180.0                                       periode:  360.0
        0      0                                       symmetry:  180.0
  0  10  20  30  40  50  60  70  80  90 100 110 120 130 140 150 160 170 180
  0   0   0   1   1   8  29  75 138 107  78  43  31  18  27  44  73 177 177
-------------------------------------------------------------------------
entry 9: open-chain ethers (s. Fig. 2, bottom row)
CANY 1  O3 1  C3(H)(H) 1  CANY
     90.0  180.0                                       periode:  360.0
       10      0                                       symmetry:  180.0
  0  10  20  30  40  50  60  70  80  90 100 110 120 130 140 150 160 170 180
  1   2   2   3   3   2   8  39  45  37  21  12   8  13  17  27 131 496 539
-------------------------------------------------------------------------
```

Fig. 3. Torsion-angle database entries for differently substituted anisoles (1–3, see Fig. 1, left), phenoxyesters (4–6, see Fig. 1, right), and open-chain alkanes (7, see Fig. 2, upper row), amines (8, Fig. 2, central row) or ethers (9, Fig. 2, bottom row). In the first row of each entry, the fragment is coded using an atom and bond-type notation similar to that in SYBYL. In the second and third row, the preferred angular values and the assigned energy ranking are given, together with the symmetry and periodicity of each individual fragment. The two bottom rows show the angular range (in 10° bins) for each torsional fragment and the angular distribution as observed in crystal structures (the complete list of database entries used in the program has been deposited as supplementary material).

The first three terms have been calculated in a previous step. The additional term E_{ac} describes the overlap penalty between fragments a and c. It is only computed if the first three terms do not exceed the predefined energy threshold. Furthermore, the spatial position of fragment a relative to c can be determined without explicitly using the coordinates of moiety B. Only the relative positions and orientations of the bonds linking a to B and B to c have to be known, together with the torsion angles set for the considered conformation. In branched substructures a-B(-c)-d (B being the central moiety), the ranking is given by contributions already determined in previous steps:

$$E_{a\text{-}B(\text{-}c)\text{-}d} = E_{a\text{-}B\text{-}c} + E_{a\text{-}B\text{-}d} + E_{c\text{-}B\text{-}d} - E_{a\text{-}B} - E_{B\text{-}c} - E_{B\text{-}d} + E_{B}$$

For those combinations that survive the ranking process as being 'favorable', coordinates in space are produced.

Structure optimization

After the combination procedure has been completed, those generated conformations which are ranked as energetically more favorable (the threshold is given by the user) can be subjected to an optimization of the open-chain torsion angles. The program TORSO, derived from the molecular mechanics program MOMO [12], uses only van der Waals [13] and torsion-angle potentials. The latter 'potentials' are derived from the statistically evaluated torsional fragments retrieved from the CSD. In the optimization procedure, the program tries to minimize unfavorable steric contacts between adjacent atoms in a molecule. A quasi-Newton method [14] is used, applying a line search technique described by Dennis and Schnabel [15]. All derivatives are calculated analytically.

Torsional potentials

In order to transform the probability distributions of crystallographically observed angles into an empirical potential, an approach described by Murray-Rust [16] has been applied. The potential value at a given angle τ is expressed as the negative logarithm of its frequency:

$$\mathrm{Pot}(\tau) = -A \ln f(\tau)$$

where $\mathrm{Pot}(\tau)$ = the potential value at angle τ, A = an adjustable parameter, and $f(\tau)$ = the frequency at τ (≥ 1). The logarithm is multiplied by an adjustable parameter (see below). The potential values are calculated in steps of 10°. In those angular ranges where no experimental data are found (frequency = 0), potential values are interpolated by a quadratic polynomial which meets a maximum set to the adjustable parameter mentioned above. This polynomial coincides with the potential defined at the boundaries of the ranges where experimental data are found. Subsequently, all potential values are shifted in order to calibrate their minimal values to zero. If more than one distribution in the database could be matched onto a probe fragment, each distribution is transformed separately. Finally, these potentials are averaged at each of the 10° steps.

To obtain a smooth, periodic and monotonic functional form, a spline function (polynomial of the order 5) is interpolated between these 10° steps, using an algorithm described by Dougherty et al. [17]. If more than one distribution in the database aligns with the torsion angle

under consideration, the probabilities from the various distributions are averaged and subsequently used in spline function fitting.

Murray-Rust calibrated the adjustable parameter to 0.63 kJ/mol [16]. In the present approach, this value has been multiplied by 10 in order to adapt the scaling of the empirical potentials to those of the van der Waals interactions. For open-chain torsion angles about single bonds, where no database information is present, the following values are assumed with equal ranking: sp^3-sp^3: –180° modulo 120°; sp^2-sp^3: –180° modulo 30°; and sp^2-sp^2: –180° modulo 90°. In the optimization procedure, periodic potentials are assigned which possess minima at these values. Optionally, the user can define which molecular portions should be included into the analysis. The remaining parts are used unchanged in their initially defined geometry.

It has to be remembered that the present approach is entirely based on empirical concepts. There is no formal connection to transform data distributions retrieved from crystal structures into potential energies. It was first mentioned by Bürgi and Dunitz [18,19] that these distributions tend to congregate in the low-lying regions of the potential energy surface that underlies changes of the structural parameters under consideration. However, the data scatter provides qualitative information about the shape of the potential energy surface, at least in its low-energy regions. In order to obtain quantitative results, the probability distributions would have to mimic a Boltzmann distribution at a certain temperature. First of all, such an analysis requires a well-distributed and statistically unbiased data set with no over-representation of certain classes of compounds. This prerequisite is hardly ever fulfilled. Furthermore, there is no straightforward way to determine for each individual crystal structure the absolute deformation energy which operates on a particular fragment in its crystal environment.

In the context of the present heuristic approach, it has to be reflected whether any description of the exact shape or absolute height of the torsional potentials beyond a qualitative estimate is of great importance. Two examples may illustrate the problems involved. In principle, the analysis of mean-square displacement amplitudes, as determined from crystal diffraction data, allows one to estimate the internal barrier to rotation (chiefly torsional) for many different kinds of fragments in the crystal environment. A study by Maverick et al. [20] shows that these barriers can vary appreciably with the crystal-field environment. 2,2',4,4',6,6'-Hexa-*t*-butylazobenzene contains chemically equivalent *t*-butyl groups that all undergo internal motion. In the solid state, these six groups are all in different environments. For example, the motions of the two chemically equivalent para-groups differ markedly in the crystal. Their barriers to rotation, estimated from the anisotropic displacement parameters and force-field calculations in the solid state (between 7 and 15 kJ/mol) differ by a factor of two. However, the corresponding potential wells show approximately the same width near their minima. Thus, a structured molecular environment has a substantial influence on intramolecular torsion barriers, here of two hydrophobic *t*-butyl groups. On the other hand, if a molecule interacts with its environment, for example through strong hydrogen bonds, the electron density distribution in the molecule will be altered. This will clearly influence the heights of barriers to internal rotation. Quantum-chemical calculations on the binding geometry of creatine bound to the enzyme creatinase demonstrated the dramatic influence of hydrogen bonding on the internal barrier to rotation [21]. Accordingly, the relative energy content of a given conformation and the height of internal torsion barriers strongly depend upon interactions with its environment. In MIMUMBA, a screening of plausible conformations adopted in a highly structured environment is attempted. It will be nearly impossible to estimate to what

extent such an 'implicitly considered' environment will alter the height of the torsional barriers. Thus, it is assumed that the number and location of torsional minima, together with the shape of the potential in the low-lying regions, are correctly approximated by the present approach. Presumably, these aspects are more important for the attempted conformational search than the absolute height and form of the torsional potentials far apart from the local minima. The objective of these derived 'potentials' in the minimization of steric strain is to keep the torsion angles as close as possible to the experimentally most frequently observed values. According to the actual scatter of these values, a particular angle is considered as 'easily' or 'difficultly' deformable.

In a final step, the obtained conformers are energetically and structurally (least-squares atom-by-atom superposition) compared, in order to reject multiply generated geometries. In addition, the program can optionally perform a cluster analysis to classify the obtained conformations into families with related geometry. This part has been adapted from MOMO [12]. The program is written in Fortran, with some subroutines in Pascal. At present, it comprises more than 37 000 lines of code.

THE PROGRAM IN OPERATION: A CASE STUDY

The different steps performed by the program are illustrated using adenosine monophosphate as a case study. The probe fragment contains a furanose ring as flexible ring portion and a side chain with four rotatable bonds. It frequently occurs (as adenosine or guanosine monophosphate) as part of cofactors (e.g., NAD, NADP, FAD, ADP, GDP, ATP) bound to protein receptors. Its conformational properties have been studied in detail [2] by comparing structural data from small-molecule crystal structures with those from ligand–enzyme structures determined by protein crystallography.

A particular conformation of the probe fragment can be defined by 21 torsion angles about the various single bonds. A principal component analysis on these angles, as observed in the small-

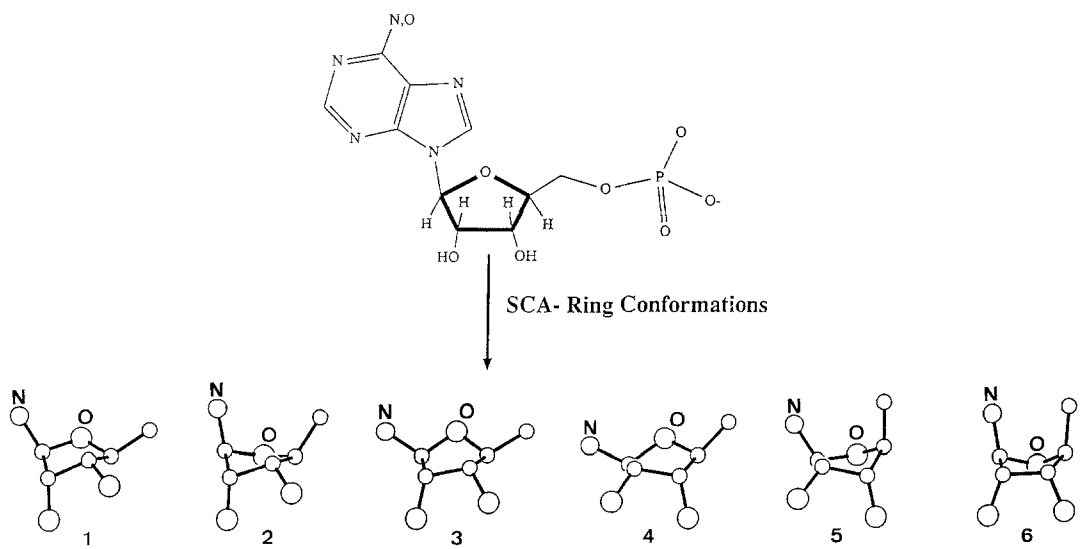


Fig. 4. Six conformations generated by the program for the flexible ribose ring in the probe fragment. (1) C3'-endo; (2) C2'-exo–C3'-endo; (3) C2'-exo; (4) C2'-endo; (5) C2'-endo–C3'-exo; (6) C3'-exo. Two of these (1 and 4) are also observed in crystal structures containing this fragment.

molecule data, reveals that the data mainly cluster about the C2'- and C3'-endo conformations of the ribose ring. The remaining variance can be ascribed to the torsion angles in the side chain. Accordingly, the four representative angles τ_{14}, τ_{16}, τ_{17} and τ_{20} have to be considered (Fig. 5).

The analysis of the ring portion with MIMUMBA reveals six low-energy conformations of the five-membered ring (Fig. 4), which can be classified as C3'-endo, C2'-exo–C3'-endo, C2'-exo, C2'-endo, C3'-exo, and C2'-endo–C3'-exo [22]. An energy ranking of 0, 1, 1, 3, 3 and 10 kJ/mol is attributed to these ring conformations. The two experimentally observed conformations are among those generated by the program.

In order to analyze the open-chain torsion angles, the program matches the four distributions given in Fig. 5. Overlapping or multiply occurring fragments, for example the threefold symmetry of the C-O-P-O fragment, is considered by superimposing the corresponding histogram with 120° periodicity. Starting structures are selected according to the values listed in Table 1, using the given energy ranking. After recomposing the individual fragments in their various conformations, only those below a limit of 30 kJ/mol are accepted for further optimization in TORSO. For this purpose, the program transforms the histograms into potentials, as shown in Fig. 6. The subse-

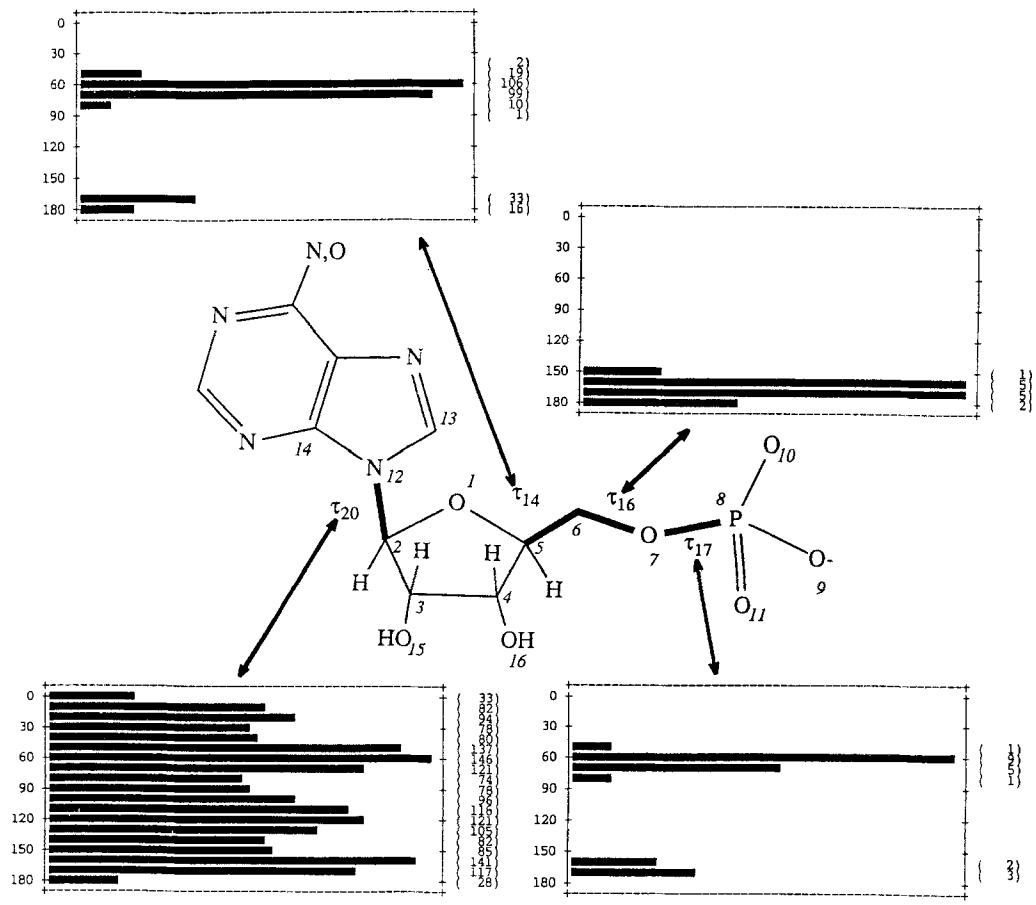


Fig. 5. Histograms of torsion-angle distributions in fragments related to the angles τ_{14}, τ_{16}, τ_{17} and τ_{20}. They are matched by the program and used after superimposing the histogram for τ_{17} according to the threefold symmetry of the OPO_3 group, with 120° periodicity to derive torsion-angle potentials (see Fig. 5).

TABLE I
TORSION-ANGLE VALUES USED IN MIMUMBA TO SELECT STARTING STRUCTURES

Torsion angle	Value
τ_{14}	−60° (0), 60° (0), 180° (5)
τ_{16}	−180° (0)
τ_{17}	−60° (2)
τ_{20}	−160° (3), −120° (0), −60° (2), −20° (0), 60° (2), 120° (0), 160° (3)

The relative energy ranking (kJ/mol) is given in parentheses.

quent optimization reveals 51 different conformations. In a subset of 19 conformers, the ribose ring displays a C2'- or C3'-endo conformation. The total analysis required 85 s on a VAX 6600, including a cluster analysis of the final conformations.

To assess the relevance of the generated conformations with respect to the geometry of the probe fragment adopted at the binding site of a protein or in a molecular crystal, the subset of 19 conformations with the C2'- or C3'-endo conformation of the ribose ring is compared to the angular distribution in the experimentally determined structures. Figure 7 summarizes the univariate statistics of τ_{14}, τ_{16}, τ_{17} and τ_{20} as observed in the small-molecule crystal structures, ligand–protein complexes, or generated by MIMUMBA. In addition, since torsion angles are accessible from protein data only with limited precision, the distances d_3 between the center of

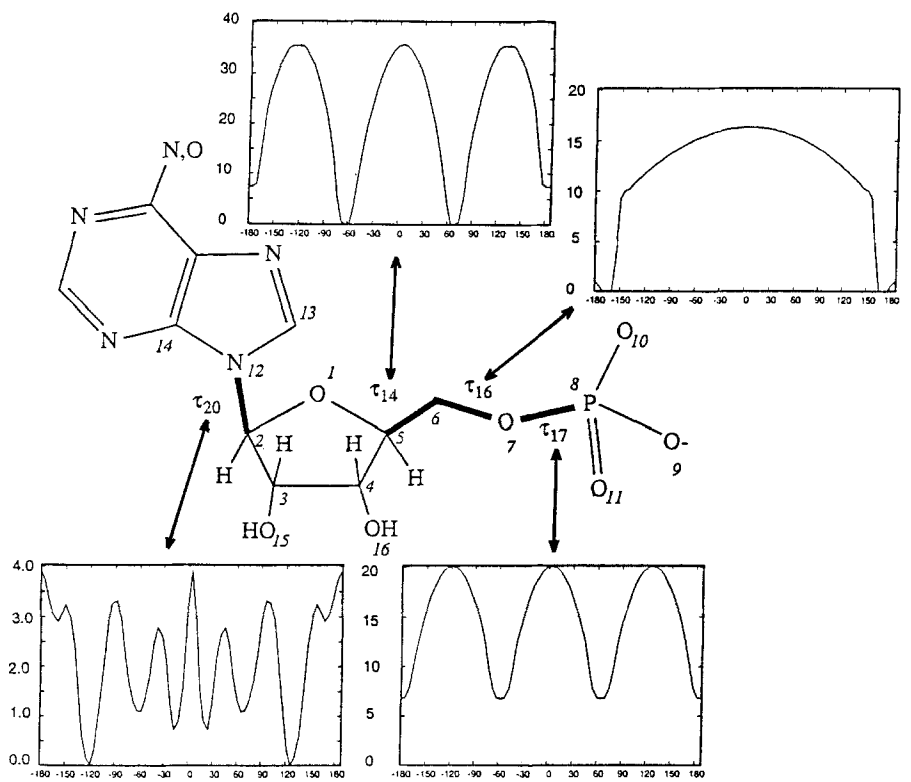


Fig. 6. Torsion-angle potentials, derived from the distributions shown in Fig. 5 and used in the optimization procedure in TORSO.

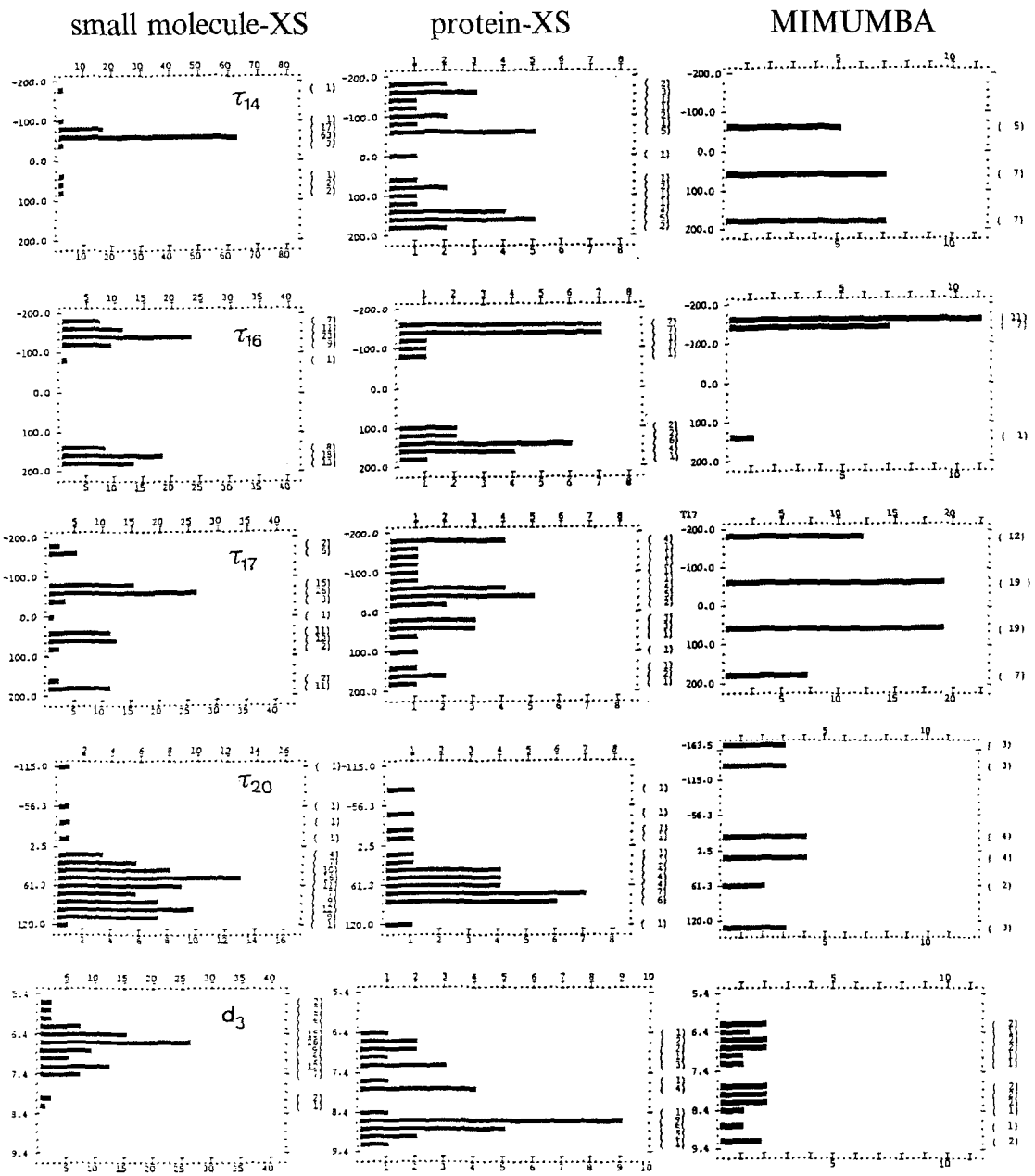


Fig. 7. Univariate statistics of the torsion angles τ_{14}, τ_{16}, τ_{17}, τ_{20} and distance d_3 (in degrees or Å; see Fig. 2 for the torsion-angle definition) in adenosine or guanosine monophosphate fragments, as found in small-molecule crystal structures (left column) or in ligands bound to proteins (central column). The right column gives the angular distributions obtained from the conformers generated by MIMUMBA (due to symmetry, τ_{17} is shown together with τ_{18} and τ_{19}).

the purine moiety and phosphorus are given.

Comparing small-molecule and protein data for τ_{16}, τ_{17} and τ_{20}, reasonable agreement is found: τ_{16} is always close to 180°, τ_{17} falls near ±60° or 180° and τ_{20} clusters between 0° and +90° (Fig. 7).

Data from protein structures are less sharply clustered, a result which is due to the limited accuracy of protein structure determinations. The angle τ_{14} shows remarkable differences between protein and small-molecule structures. Whereas in the low-molecular-weight compounds the O-C-C-O-fragment mostly occurs in a g^--conformation, for the protein environment a broad distribution is observed (g^+, g^-, t). In agreement with this observation, the distance d_3 is found to be about 6.5 Å in most small-molecule structures (only in 3 of the 90 examples $d_3 > 7.5$ Å). In the case of the proteins, this distance clusters at 6.6, 7.8 and 8.8 Å. In the crystal structures, the molecules preferentially adopt a compact folded conformation, whereas in the protein environment both geometries, i.e., a compact and an extended one, are observed.

A comparison with the conformational multiplicity generated by MIMUMBA shows that, with 19 conformations, the program has explored most of the areas in conformational space that are also found experimentally. Only in the case of τ_{20} do the generated structures spread over a slightly larger range than the experimentally observed ones. With respect to τ_{14}, the generated data comprises both the two gauche and the trans conformations.

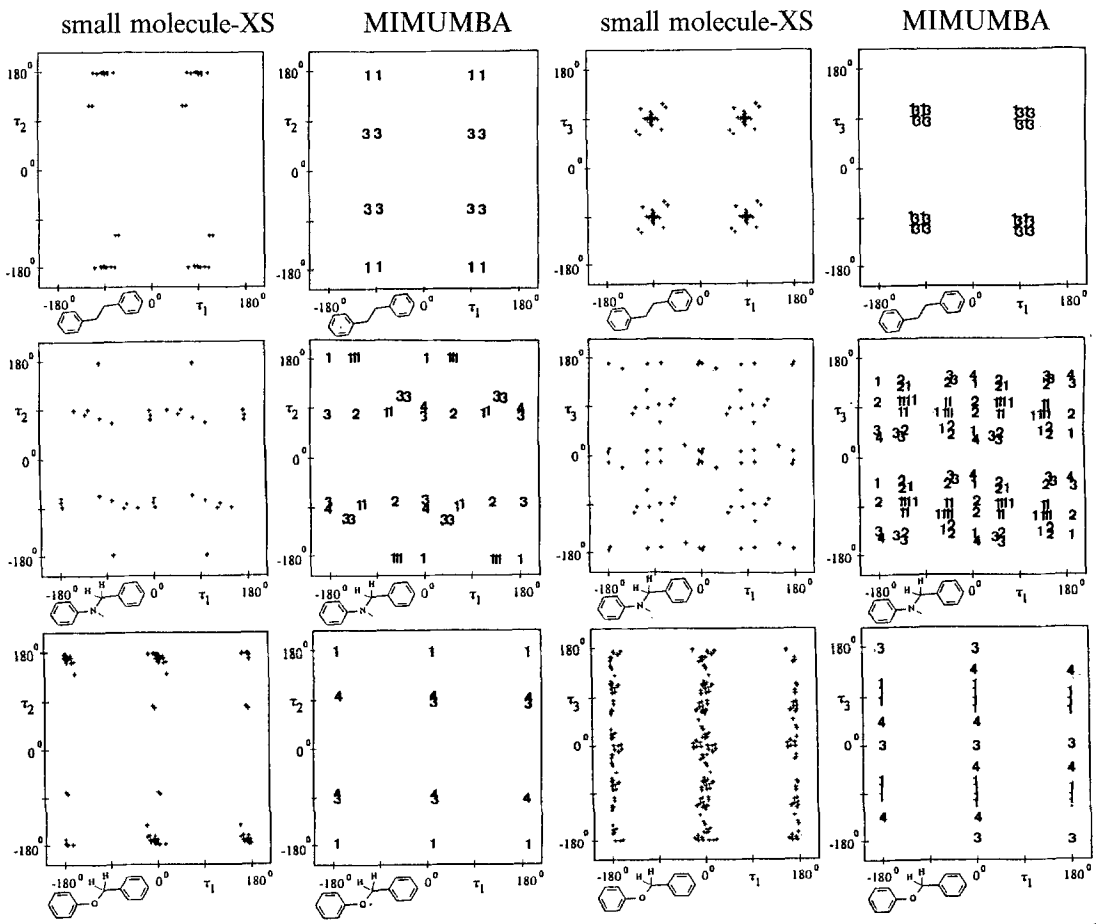


Fig. 8. Comparison of the data scatter in conformation space (torsion angles τ_1, τ_2, τ_3) for crystal structures (columns 1 and 3) containing the probe fragment Ph-X-Y-Ph (X-Y= CH_2-CH_2, NH-CH_2 or O-CH_2) and for conformers of Ph-X-Y-Ph generated by MIMUMBA (columns 2 and 4).

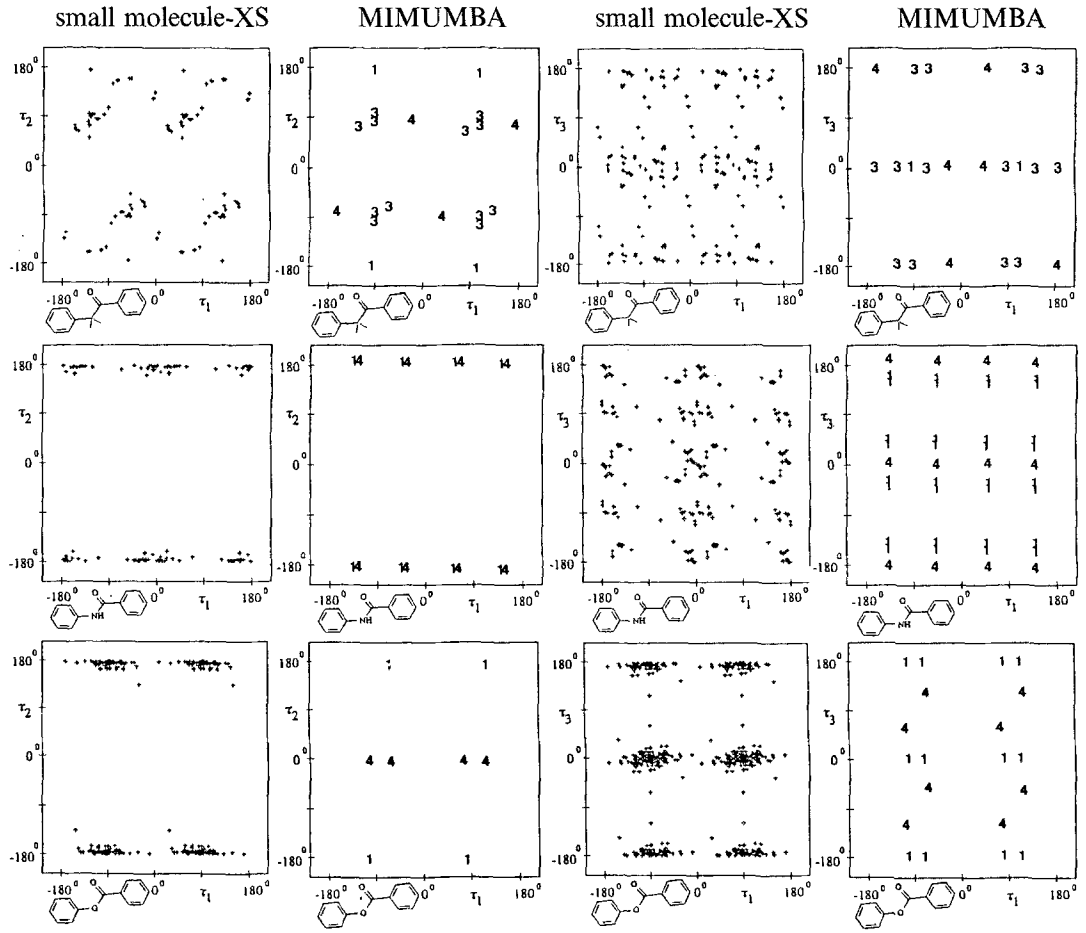


Fig. 9. Comparison of the data scatter in conformation space (torsion angles τ_1, τ_2, τ_3) for crystal structures (columns 1 and 3) containing the probe fragment Ph-X-Y-Ph (X-Y= CH_2-(C=O), NH-(C=O) or O-(C=O)) and for conformers of Ph-X-Y-Ph generated by MIMUMBA (columns 2 and 4).

ASSESSMENT OF THE COMPLETENESS OF THE CONFORMATIONAL ANALYSIS

The significance of an empirical approach can only be shown by applications that produce results similar to experimental evidence. The completeness of a conformational analysis and the quality of the produced conformers are very difficult to assess. The structure correlation method [1,18,23] is one possible approach to map conformation space of a particular molecular fragment. This approach is entirely based on experimental data. The properties of Ph-X-Y-Ph fragments have been studied by this technique [24]. Crystallographic data of molecules containing the probe fragment Ph-X-Y-Ph, where X-Y is -CH_2-CH_2-, -NH-CH_2-, -O-CH_2-, -CH_2-(C=O)-, -NH-(C=O)-, -O-(C=O)-, -CH_2-SO_2-, -NH-SO_2- and -O-SO_2-, have been retrieved from the Cambridge File (to enlarge the data set, structures with N in 3-, 4- and 5-position of the phenyl rings were considered as well). To avoid a fortuitous distribution of structural entries in conformation space, which depends on an arbitrary labeling of the atoms in each fragment, all symmetry-related arrange-

ments have been generated (τ_1,τ_2,τ_3; $\tau_1+180°,\tau_2,\tau_3$; $\tau_1,\tau_2,\tau_3+180°$; $\tau_1+180°,\tau_2,\tau_3+180°$, combined with $-\tau_1,-\tau_2,-\tau_3$; for X=Y in addition $\tau_1=\tau_3$).

For comparison, the conformation space of Ph-X-Y-Ph has been explored using MIMUMBA. Three different torsion angles have to be considered: a side chain -X-Y at a phenyl ring, an open-chain fragment -X-Y- between two carbon atoms, and a side chain -Y-X at a phenyl ring. The program matches distributions corresponding to these three torsion angles in its database and generates appropriate starting structures. After combining the individual fragments in their various conformations and optimizing those of the composed structures with an energy ranking below 30 kJ, distributions in conformation space are obtained that are summarized in Figs. 8–10. It has to be pointed out that the structure correlation approach evaluates properties of the entire fragment as found in crystal structures. In contrast, the results obtained by MIMUMBA are based on univariate statistics, derived for each considered torsion angle separately from broad samples of crystal data. In general, these samples contain structures showing only one of the three

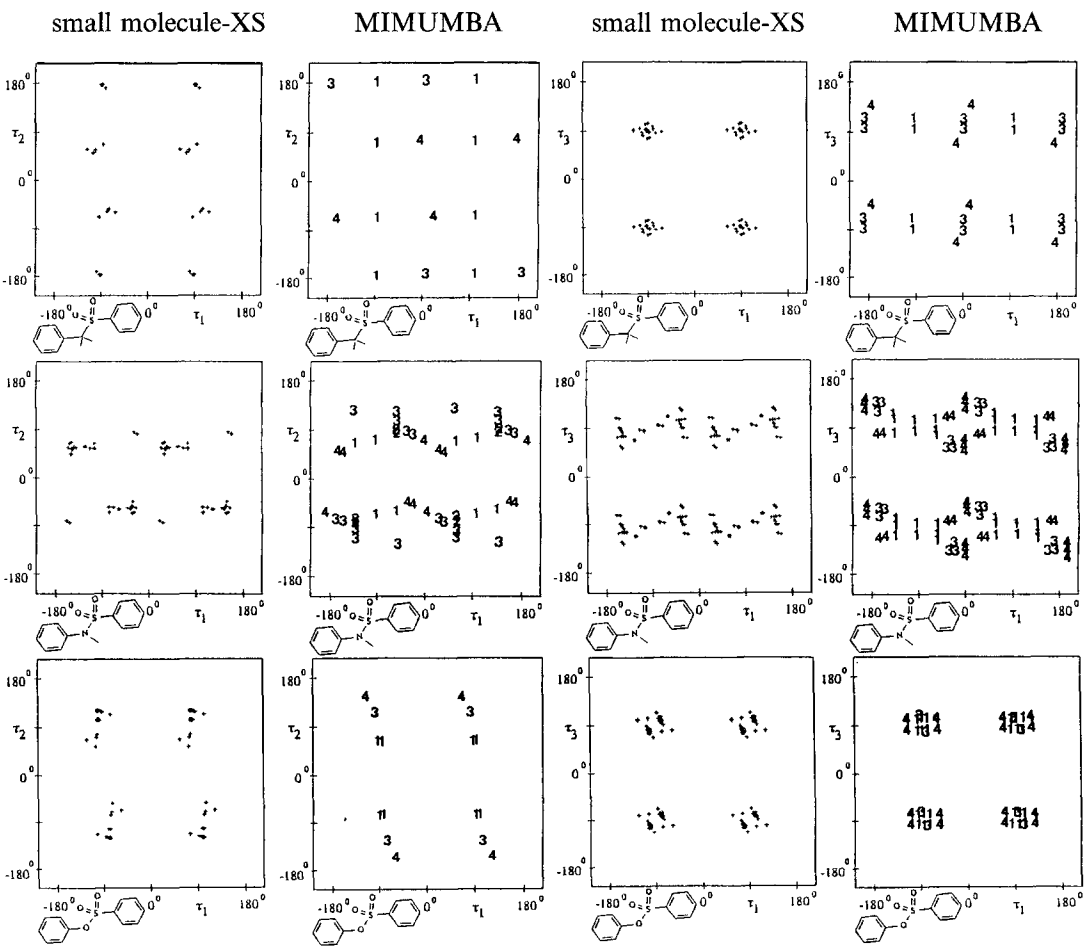


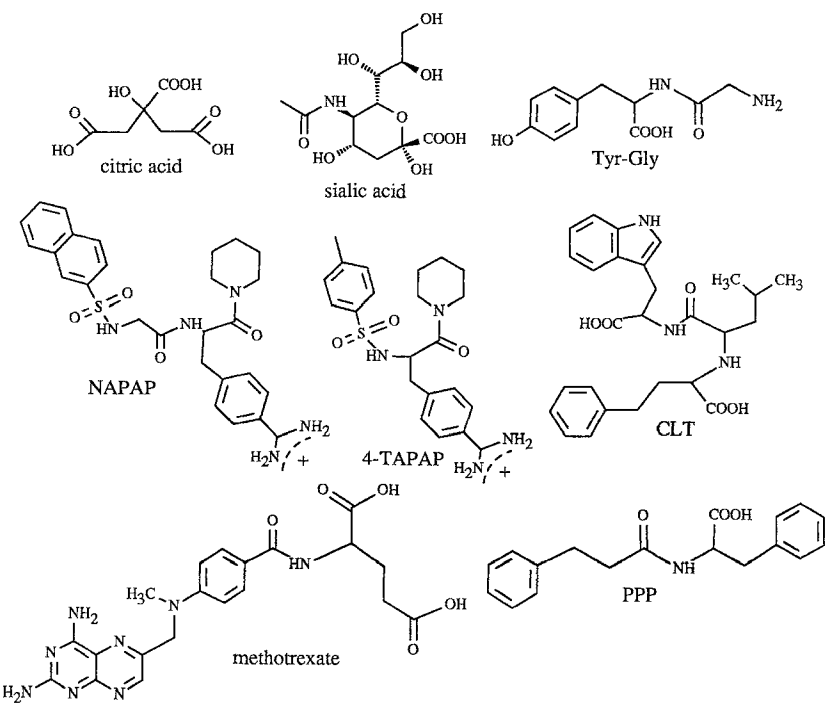
Fig. 10. Comparison of the data scatter in conformation space (torsion angles τ_1, τ_2, τ_3) for crystal structures (columns 1 and 3) containing the probe fragment Ph-X-Y-Ph (X-Y= CH_2-SO_2, NH-SO_2 or O-SO_2) and for conformers of Ph-X-Y-Ph generated by MIMUMBA (columns 2 and 4).

angles of interest in Ph-X-Y-Ph. Any restricted occupation in conformation space is obtained by MIMUMBA through a correct consideration of the van der Waals interactions in the composed conformers.

The energy ranking attributed by the program to the various conformers is indicated in the figures. The structures are grouped into four classes according to a ranking being low (=1), medium (=2 and =3), or high (=4) in energy. In most examples, conformation space is similarly mapped by the correlation method and by MIMUMBA. In some of the cases where a broad cluster of experimental data is observed, the program places one particular conformer approximately into the center of this cluster (e.g. esters; see Fig. 9, bottom row). Interestingly, conformations generated by the automatic procedure in areas of conformational space where no experimental data are found are ranked as energetically less favorable. For example, in the case of esters (X-Y: O-(C=O)) also cis-configurated esters ($\tau_2 = 0°$) are produced; however, they are classified as high-energy conformers. For sulfones Ph-CH_2-SO_2-Ph (Fig. 10, top row), conformers are generated at $\tau_1 = 0°$, $\pm 180°$ that are ranked as energetically less favorable. In summary, the obtained correspondences give some confidence that the program efficiently maps conformation space and that the van der Waals interactions are appropriately considered.

GENERATION OF LIGAND CONFORMATIONS

Finally, we examined whether, for a given protein ligand, the program actually generates a set of conformations that contains a geometry structurally close to the geometry adopted at the



Scheme 1. Ligands used to compare MIMUMBA-generated conformations with experimentally found ones.

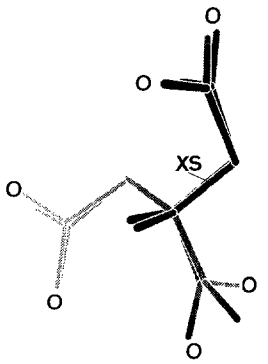


Fig. 11. Superposition of the conformation of citric acid (indicated by XS, hydrogen atoms omitted) adopted at the binding site of citrate synthase (3CTS) with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.32 Å, ranked as no. 1 out of 14). Plots were obtained with the program SHADEMOL [55].

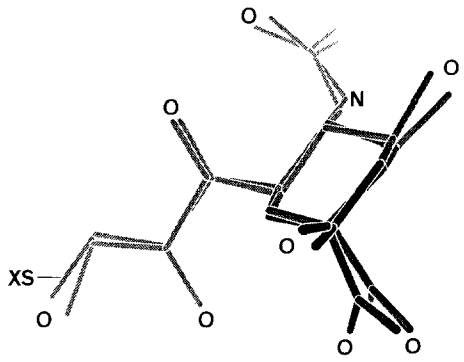


Fig. 12. Superposition of the conformation of sialic acid (indicated by XS, hydrogen atoms omitted) adopted at the binding site of hemagglutinin (4HMG) with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.39 Å, ranked as no. 11 out of 14).

binding site. The conformation of the ligand was retrieved from a ligand–protein complex as stored in the Brookhaven File [25] (references to the structures are given by their PDB codes). After assigning appropriate atom and bond types (according to the notation in SYBYL [10]) and adding hydrogen atoms in standard geometry, the structures were optimized with the MaxiMin force field [26] to obtain consistent bond lengths and angles. The chemical formulae of the various ligands used are shown in Scheme 1. The minimized structures were taken as input for conformational analysis. The conformers, generated by MIMUMBA within a predefined energy threshold (see 'ranking'), were fitted onto the references determined by protein crystallography (least-squares atom-by-atom superposition). The obtained rms deviations are given. In this context, we have to keep in mind that the experimental error for the atomic positions in protein X-ray structures is estimated to be about 0.4 Å in space [27].

As a first example, the conformation of citric acid in citrate synthase [28] (PDB code: 3CTS), shown in Fig. 11, is considered. This ligand contains five rotatable bonds. For a systematic rotation about each of these bonds in steps of 30°, more than 200 000 geometries would have to be considered. A study with MIMUMBA allowed us to find 14 conformers ranked in energy below the threshold value of 30 kJ. The lowest energy structure happens to fall close to that present at the binding site (rms = 0.32 Å, see Fig. 11).

Sialic acid contains one flexible pyranose ring. This molecule binds to hemagglutinin and has been structurally characterized at the binding site [29] (4HMG). MIMUMBA generates a set of 14 conformers, considering the six-membered ring and all side chains as flexible. The geometry observed at the binding site is closely represented by one conformer (rms = 0.39 Å, Fig. 12), ranked in the list as no. 11.

The binding geometry of the dipeptide Tyr-Gly in carboxypeptidase A (3CPA) has been determined [30]. This ligand has five rotatable bonds if the amide bond is considered rigid. A conformational search with MIMUMBA produced 60 different conformers, of which no. 6 resembles the binding geometry reasonably well (rms = 0.52 Å, Fig. 13).

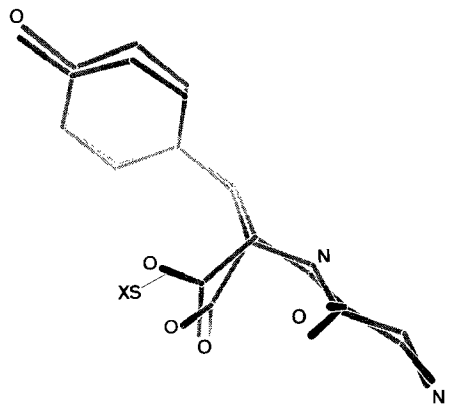


Fig. 13. Superposition of the conformation of the dipeptide Tyr-Gly (indicated by XS, hydrogen atoms omitted) adopted at the binding site of carboxypeptidase A (3CPA) with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.52 Å, ranked as no. 6 out of 60).

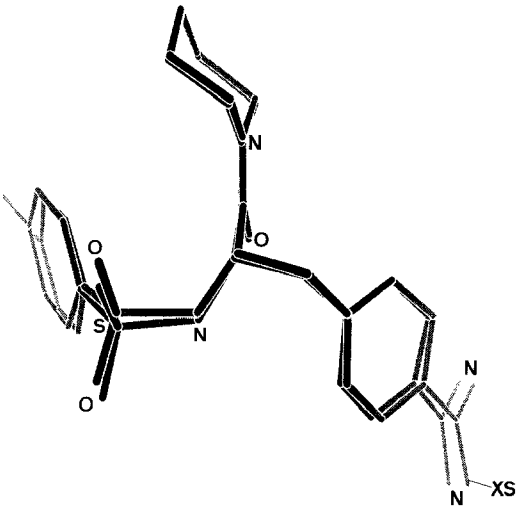


Fig. 15. Superposition of the conformation of 4-TAPAP (indicated by XS, hydrogen atoms omitted) adopted at the binding site of thrombin with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.35 Å, ranked as no. 43 out of 144).

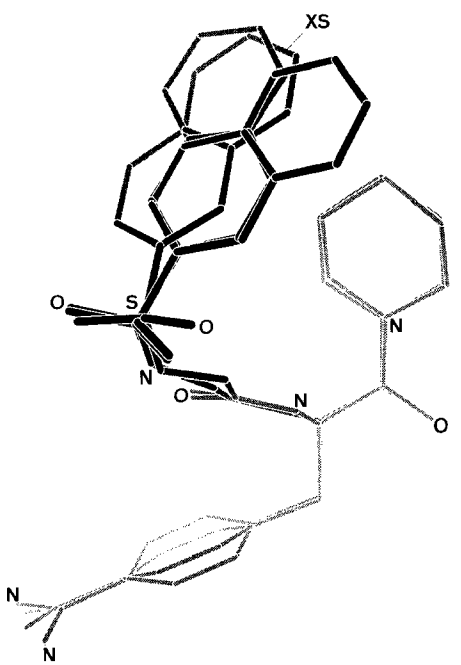


Fig. 14. Superposition of the conformation of NAPAP (indicated by XS, hydrogen atoms omitted) adopted at the binding site of thrombin [31] with two conformations generated by MIMUMBA. One of these falls close to the experimentally observed geometry (rms = 0.81 Å, ranked as no. 5 out of 89). The additional conformer (no. 8) differs in the orientation of the naphthyl ring. Some experimental evidence for the presence of this second conformer at the binding site of thrombin is indicated by a disorder reported for the naphthyl moiety in one of the structure determinations [32].

NAPAP is an inhibitor of thrombin. Two structure determinations of this ligand–enzyme complex have been reported. Whereas Brandstetter et al. [31] found a single conformation at the binding site, Banner and Hadvary [32] report a disorder. Two conformational orientations of the naphthyl ring are assumed. In a set of 89 conformers generated by MIMUMBA, two conformers are found that are close to the geometry at the binding site (no. 5, rms = 0.81 Å, and no. 8 with respect to the protein complex [32], Fig. 14). They differ in the torsion angle about the bond toward the naphthyl moiety, thus representing both conformers reported by Banner and Hadvary. In the analysis, all bonds apart from the peptide bond and the bond toward the amidino group were considered flexible. Another inhibitor of thrombin is 4-TAPAP [31]. In this ligand six rotatable bonds have to be considered. Similar to the analysis of NAPAP, the benzamidino group and the peptide bond with the piperidine ring were treated as rigid fragments. The program generated 144 conformers within the given energy-ranking threshold. The one ranked as no. 43 in the list is close to the conformation reported by crystal structure analysis (rms = 0.35 Å, Fig. 15).

A more complex example is methotrexate (MTX, Fig. 16), an inhibitor of dihydrofolate reductase (DHFR). This ligand contains nine rotatable bonds if the peptide bond is fixed in the trans geometry. In the analysis, the torsion angles toward the terminal carboxy groups were restricted to 180°. Regarding the remaining angles, a run which generated 445 conformations contains a structure (ranked as no. 242) closely related to the binding-site geometry of MTX in *L. casei* DHFR [33] (3DFR, rms = 0.76 Å, Fig. 16). In contrast, in the enzyme from *E. coli*, the ligand adopts a gauche conformation for the terminal CH_2-CH_2-COOH chain. Since a gauche conformation for this torsional fragment is ranked higher in energy than a trans arrangement,

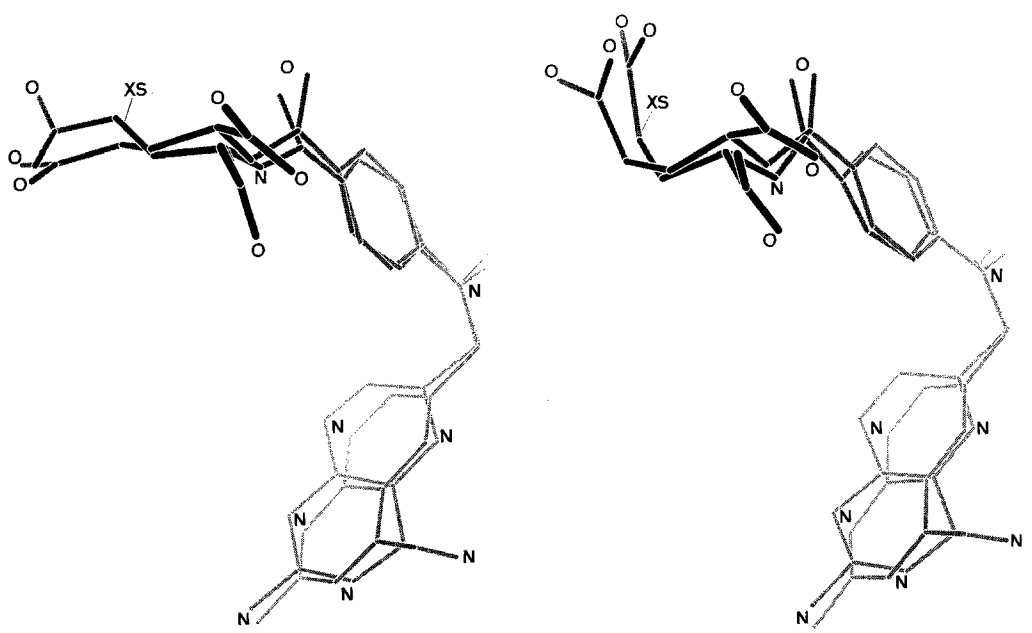


Fig. 16. Superposition of the conformation of methotrexate (indicated by XS, hydrogen atoms omitted) adopted at the binding site of two different species of dihydrofolate reductase (left: *E. coli* DHFR, 4DFR; right: *L. casei* DHFR, 3DFR) with conformations generated by MIMUMBA. These fall close to the experimentally observed geometries (3DFR: rms = 0.76 Å, ranked as no. 242 out of 445; 4DFR: rms = 0.77 Å, ranked as no. 643 out of 860).

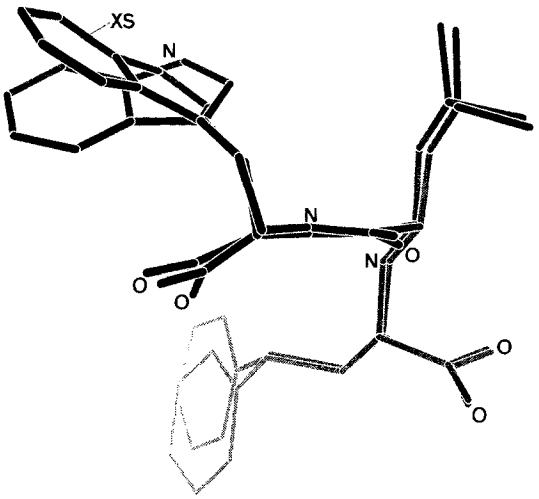


Fig. 17. Superposition of the conformation of CLT (indicated by XS, hydrogen atoms omitted) adopted at the binding site of thermolysin (1TMN) with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.82 Å, ranked as no. 28 out of 58).

a conformation similar to the binding-site geometry in *E. coli* DHFR is only found in a larger data set of generated conformations. Allowing a higher cutoff value for the energy ranking, a set of 860 conformations is obtained which now also contains the expected gauche conformation (no. 643) in the side chain (4DFR, rms = 0.77 Å, Fig. 16).

A further example of similar complexity is *N*-(1-carboxy-3-phenylpropyl)-L-leucyl-L-tryptophan (CLT), an inhibitor of thermolysin [34] (1TMN). If the peptide bond, the isobutyl group and the two carboxy groups are excluded from the analysis, the inhibitor contains nine rotatable bonds. In a run which generated 58 favorable conformations, the arrangement shown in Fig. 17 (ranked as no. 28, rms = 0.82 Å) is structurally similar to the conformation found in the enzyme-binding pocket. Another inhibitor of thermolysin is β-phenylpropionyl-L-phenylalanine (PPP) [35]. This extended molecule contains seven rotatable bonds if the peptide bond is excluded from the analysis. A run which generates 89 individual conformers of PPP includes one conformer (ranked as no. 21) which resembles that adopted at the binding site (rms = 0.62 Å, Fig. 18).

Apart from methotrexate, for most of the ligands considered a relevant conformation is found in a set of less than 100 conformers. However, the MTX example also indicates some limitations

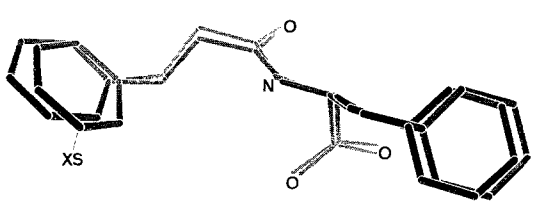


Fig. 18. Superposition of the conformation of PPP (indicated by XS, hydrogen atoms omitted) adopted at the binding site of thermolysin [35] with a conformation generated by MIMUMBA. The latter conformer falls close to the experimentally observed geometry (rms = 0.62 Å, ranked as no. 21 out of 89).

of the approach. For rather large ligands, showing one or several torsion angles in less favorable conformations, a relevant geometry can only be expected in an extended set of generated conformers.

SUMMARY AND CONCLUSIONS

Several approaches toward conformational analysis have been described in the literature (for recent reviews, see Refs. 36 and 37). Most molecular mechanics and quantum-chemical programs provide a driver option for scanning torsion angles. Considering several torsional degrees of freedom at a time by this approach is normally prohibitive, because of its enormous computational requirements. Molecular dynamics at elevated temperature or stochastic Monte Carlo methods have been applied to conformational searching [38–40]. Distance geometry is an alternative approach for probing conformation space [41]. However, the latter techniques produce rather distorted starting structures, thus elaborate geometry optimizations are required to obtain reasonable geometries. Some comparative studies have been performed to elucidate the efficiency of the various approaches [42,43].

Some techniques, especially tailored to search for biologically relevant conformers have been reported. The most popular approach in the field is the systematic search; however, due to the overwhelming amount of produced geometries it is rather unefficient. In combination with the active analog approach developed by Marshall et al. [44,45], its efficiency can be improved. A data set of structurally diverse and conformationally complementary structures is required and some assumptions about common pharmacophoric groups have to be defined. Under these conditions, the accessibility of conformation space is sufficiently constrained to reveal only a limited number of possible conformers [46]. Another approach described in the literature [47] is implemented into the program ALADDIN. Similar to the previous method, this program tries to test conformational requirements for bioactivity within a set of given drug molecules.

All methods listed so far are either computationally very intensive or produce an overwhelming amount of data. This clearly limits their broad application to routine studies in drug design. A knowledge-based approach toward conformational analysis is presented in this paper. Other knowledge-based approaches have previously been described in the literature. The Chem-X 3D Builder [48], WIZARD and COBRA [49–52] are based on conformational templates and sets of rules derived from molecular mechanics. The program AIMB [53] uses, similar to the present approach, crystallographic data as input. It generates multiple conformations of molecules by assembling the largest and most appropriate fragments in a knowledge base, composed of crystal structures.

We based our approach on the empirical observation that conformational preferences and, accordingly, torsion-angle libraries of molecular fragments retrieved from large sets of small-molecule crystal structures indicate also relevant geometries of low-molecular-weight ligands at the binding site of proteins. Using this information, we developed a fast and efficient method to generate ligand conformations (about 1–5 s per conformer on a Silicon Graphics RS3000 or 4000 workstation [54]). The produced conformers possess some probability to represent geometries that are likely to resemble those actually adopted at the binding site of a protein. Sets of some 10 to 100 conformers generated for ligands of known active-site geometry (from protein crystallography) contain conformers that closely approach the experimentally observed arrangement. This

aspect is especially important in those cases where the 3D structure of the protein receptor is not known. It can be hoped that already a limited number of conformers allows one to derive reasonable estimates for the biologically active conformation.

Local minima obtained by the program fall close to geometries actually found at the binding site of proteins. However, highly distorted geometries adopted by substrate molecules on their way toward the transition state of, for example, an enzyme reaction, cannot be handled. The approach does not explicitly consider intramolecular hydrogen bonds. Nevertheless, if conformational properties of a torsional fragment (for example ϕ and ψ angles in peptide-like molecules) were retrieved from structures that frequently occur with intramolecular hydrogen bonds in the crystal (e.g. β-turns), MIMUMBA will generate conformers showing this geometry.

ACKNOWLEDGEMENTS

The authors are grateful to W. Berger, U. Hesse and F. Weber for their help in data retrieval and evaluation.

REFERENCES

1 Bürgi, H.B. and Dunitz, J.D., Structure Correlation, Vol. 1 and 2, VCH, Weinheim, 1994.
2 Klebe, G., In Bürgi, H.B. and Dunitz, J.D. (Eds.) Structure Correlation, Vol. 2, VCH, Weinheim, 1994, pp. 543–603.
3 Ricketts, E.M., Bradshaw, J., Hann, M., Hayes, F., Tanna, N. and Ricketts, D.M., J. Chem. Inf. Comput. Sci., 33 (1993) 905.
4 Pearlman, R.S., CDA News, 2 (1987) 1.
5 Gasteiger, J., Rudolph, C. and Sadowski, J., Tetrahedron Comput. Methodol., 3 (1990) 537.
6 DeClerq, P.J., Hoflack, J. and Cauwbergh, S., QCPE Program No. QCMP079, Bloomington, IN.
7 Allen, F.H. and Kennard, O., Acc. Chem. Res., 16 (1983) 146.
8 Allen, F.H., Kennard, O., Watson, D.G., Brammer, L., Orpen, A.G. and Taylor, R., J. Chem. Soc., Perkin Trans. II, (1987) S1.
9 Bürgi, H.B. and Dunitz, J.D., In Bürgi, H.B. and Dunitz, J.D. (Eds.) Structure Correlation, Vol. 1, VCH, Weinheim, 1994, pp. 23–70.
10 SYBYL Molecular Modeling System (Version 5.40), Tripos Associates, St. Louis, MO.
11 On request, the supplementary material can be made available through electronic mail provided an appropriate e-mail address is delivered.
12 Beck, H. and Egert, E., force-field program MOMO, University of Göttingen, Göttingen, 1988.
13 Smith, A.E. and Lindner, H., J. Comput.-Aided Mol. Design, 5 (1991) 235.
14 Davidon, W.C., Math. Programming, 9 (1975) 1.
15 Dennis, J.E. and Schnabel, R.B., Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
16 Murray-Rust, P., In Griffin, J.F. and Duax, W.L. (Eds.) Molecular Structure and Biological Activity, Elsevier, New York, NY, 1982, pp. 117–133.
17 Dougherty, R.L., Edelman, A. and Hyman, J.M., Math. Comput., 52 (1989) 471.
18 Bürgi, H.B. and Dunitz, J.D., Acc. Chem. Res., 16 (1983) 153.
19 Bürgi, H.B. and Dunitz, J.D., Acta Crystallogr., B44 (1988) 445.
20 Maverick, E., Mirsky, K., Knobler, C.B. and Trueblood, K.N., Acta Crystallogr., B47 (1991) 272.
21 Klebe, G., In Jeffrey, G.A. and Piniella, J.F. (Eds.) The Application of Charge Density Research to Chemistry and Drug Design, Plenum, New York, NY, 1991, pp. 287–318.
22 Saenger, W., Angew. Chem., 85 (1973) 680; Angew. Chem., Int. Ed. Engl., 12 (1973) 591.
23 Schweizer, B., In Bürgi, H.B. and Dunitz, J.D. (Eds.) Structure Correlation, Vol. 1, VCH, Weinheim, 1994, pp. 369–404.
24 Klebe, G., J. Struct. Chem. (THEOCHEM), 308 (1994) 53.

25 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rogers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
26 Clark, M., Cramer III, R.D. and Van Opdenbosch, N., J. Comput. Chem., 10 (1989) 982.
27 Dauber-Osguthorpe, P., Roberts, V.A., Osguthorpe, D.J., Wolff, J., Genest, M. and Hagler, A.T., Proteins, 4 (1988) 31.
28 Remington, S., Wiegand, G. and Huber, R., J. Mol. Biol., 158 (1982) 111.
29 Sauter, N.K., Glick, G.D., Crowther, R.L., Park, S.J., Eisen, M.B., Skehel, J.J., Knowles, J.R. and Wiley, D.C., Proc. Natl. Acad. Sci. USA, 89 (1992) 324.
30 Christianson, D.W. and Lipscomb, W.N., Proc. Natl. Acad. Sci. USA, 83 (1986) 7568.
31 Brandstetter, H., Turk, D., Hoeffken, H.W., Grosse, D., Stürzebecher, J., Martin, P.D., Edwards, B.F.P. and Bode, W., J. Mol. Biol., 226 (1992) 1085.
32 Banner, D.W. and Hadvary, P., J. Biol. Chem., 266 (1991) 20085.
33 Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. and Kraut, J., J. Biol. Chem., 257 (1982) 13650.
34 Monzingo, A.F. and Matthews, B.W., Biochemistry, 23 (1984) 5724.
35 Kester, W.R. and Matthews, B.W., Biochemistry, 16 (1977) 2506.
36 Marshall, G.R., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, 1993, pp. 80–116.
37 Leach, A.R., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Reviews in Computational Chemistry II, VCH, New York, NY, 1991, pp. 1–55.
38 Hagler, A.T., Osguthorpe, D.J., Dauber-Osguthorpe, P. and Hemple, J.C., Science, 227 (1985) 1309.
39 Ferguson, D.M. and Raber, D.J., J. Am. Chem. Soc., 111 (1989) 4371.
40 Saunders, M., J. Am. Chem. Soc., 109 (1987) 3150.
41 Crippen, G.M. and Havel, T.F., Distance Geometry and Molecular Conformation, Research Studies Press, Wiley, New York, NY, 1988.
42 Saunders, M., Houk, K.N., Wu, Y.D., Still, W.C., Lipton, M., Chang, G. and Guida, W.C., J. Am. Chem. Soc., 112 (1990) 1419.
43 Böhm, H.J., Klebe, G., Lorenz, T., Mietzner, T. and Siggel, L., J. Comput. Chem., 11 (1990) 1021.
44 Marshall, G., Annu. Rev. Pharmacol. Toxicol., 27 (1987) 193.
45 Marshall, G.R., Barry, C.D., Bosshard, H.E., Dammkoehler, R.D. and Dunn, D.A., In Olson, E.C. and Christoffersen, R.E. (Eds.) Computer-Assisted Drug Design, Vol. 112, American Chemical Society, Washington, DC, 1979, pp. 205–226.
46 Mayer, D., Naylor, C.B., Motoc, I. and Marshall, G.R., J. Comput.-Aided Mol. Design, 1 (1987) 3.
47 Van Drie, J.H., Weininger, D. and Martin, Y.C., J. Comput.-Aided Mol. Design, 3 (1989) 225.
48 Davies, K. and Upton, R., Tetrahedron Comput. Methodol., 3 (1990) 665.
49 Dolata, D.P. and Carter, R.E., J. Chem. Inf. Comput. Sci., 27 (1987) 36.
50 Dolata, D.P., Leach, A.R. and Prout, K., J. Comput.-Aided Mol. Design, 1 (1987) 73.
51 Leach, A.R., Prout, K. and Dolata, D.P., J. Comput.-Aided Mol. Design, 2 (1988) 107.
52 Leach, A.R. and Prout, K., J. Comput. Chem., 11 (1990) 1193.
53 Wipke, W.T. and Hahn, M.A., Tetrahedron Comput. Methodol., 1 (1988) 141.
54 As an example: on a Silicon Graphics Indigo RS4000 the average time to generate a conformer of NAPAP (Fig. 14) was 3.2 s; for the smaller inhibitor 4-TAPAP (Fig. 15) 1.8 s per conformer were required.
55 Hahn, M. and Wipke, W.T., Tetrahedron Comput. Methodol., 1 (1988) 81.