A Systematic Docking Approach. Application to the α-Cyclodextrin/Phenyl-Ethanol Complex

V. TRAN, M. M. DELAGE, and A. BULÉON
INRA Nantes, Laboratoire de Physico-Chimie des Macromolécules rue de la Géraudière.
BP 527-44026 Nantes Cedex 03, France.

Abstract. A new approach for systematic docking is applied to the structure of the α-cyclodextrin/phenyl-ethanol complex. This methodology includes systematic scanning of the possible guest positions, clustering of low energy structures into families and final refinement using molecular mechanics. The clustering was performed on internal parameters of the complex by a program named PROXIM based on a very simple proximity criterion. This program organized nearly 30 000 structures into about 100 families. Thirty conformations have been considered (10 and 20 for the complexation on the primary and secondary face respectively), the two forms of complexation encountered in the crystal packing yield the lowest energy combination.

Key words. Molecular modelling, docking, clustering, complexation, cyclodextrins.

1. Introduction

Cyclodextrins (CD) are cyclic oligosaccharides consisting of six (α), seven (β) or eight (γ) glucose units covalently linked by α(1–4) bonds. They admit various guest molecules into their hydrophobic toroidal cavity to form inclusion complexes. Because of this, cyclodextrins are extensively studied as model compounds for enzyme–substrate reactions and donor–acceptor binding [1–4] and have wide applications in the food and pharmaceutical industries [5–7]. The driving forces for complexation are attributed to different types of interactions: dipole–dipole interaction [8], hydrogen bonding [9], van der Waals interaction [10, 11], hydrophobic interaction [12] or release of high-energy cavity water [13].

The structures of many CD complexes have been resolved by means of single crystal diffraction techniques [14] yielding detailed information on the structure of the host lattice and the guest–host interactions, including the hydrogen bonds involved and the water molecules present in the structure. Solid state NMR and neutron diffraction have also been used to follow the structural modifications which accompany changes in the level of hydration [15–17]. In solution, the formation of CD complexes can be analyzed with various experimental techniques such as visible and UV spectroscopies, circular dichroism, ^{1}H and ^{13}C NMR spectroscopy [7, 18, 19].

Molecular modelling can be used to predict the interactions between CD and guest molecules and to propose three dimensional structures for the CD complexes when crystallization and X-ray measurements are unavailable. Molecular graphics [20] and molecular mechanics have been used to calculate the van der Waals and the electrostatic contributions to the total interaction energy [21, 23] and to define

the orientation of the guest in the host cavity. Other studies using quantum mechanics and molecular orbital approaches provide information about the contribution of the dipole–dipole interaction in the formation of the complexes [8, 24]. Molecular dynamics simulations have also been used to predict the structures of CD complexes, especially to take account of the flexibility of the host and the positions of water molecules and hydrogen bonds [25–27]. Stochastic methods, such as dynamical simulated annealing [28], can also be used to study host–guest interactions.

Exhaustive systematic searching of conformations is the most appropriate for the study of interactions between two molecules (docking) when the number of parameters is limited. Nevertheless, only a few techniques of systematic docking have been described [29–31]. Typically, a guest is manually positioned in the host's functional zone, with starting geometries derived from crystallographic information. Further molecular mechanics minimizations lead to an acceptable interaction geometry in terms of energy, but this result frequently corresponds to the conformation of minimum energy closest to the starting point (i.e. the initial manual positioning).

This paper describes a methodology involving a systematic scanning of the parameters defining the relative positions of guest and host molecules. This strategy consumes more time but accesses all possible structures. Because fairly non-restrictive criteria are used, a great many structures are seen as plausible. Therefore, two major problems must be solved: the selection of the most probable results and their clustering into the different possible structures for the studied complex. This method has been tested on the αCD/phenyl-ethanol (PE) complex for which a structure has been determined by X-ray crystallography [32]. It could also be suitable for predicting the structures of unknown complexes of CD or amylose.

2. Methods

The proposed strategy includes three main steps:

1. Systematic scanning of the positioning parameters with exclusion of improbable structures,
2. Clustering of the remaining structures into families and description of the families with 'synthetic' parameters,
3. Final refinement of the best structures representing the different families.

This strategy was carried out using the SYBYL package [33] with the standard options, except for the clustering aspect, which was solved with an additional algorithm named PROXIM. This program could be easily added or integrated to other packages. Calculations were performed on either a Micro Vax 3100 or Vaxstations (3500, 3200) and the structures were viewed on an Evans & Sutherland PS 350 graphics terminal.

2.1. SYSTEMATIC SCANNING

The crystal structure of αCD/PE [32] was obtained from the Cambridge Crystallographic Database. The stacking of αCD and PE molecules along the a axis of the unit cell is shown in Figure 1.
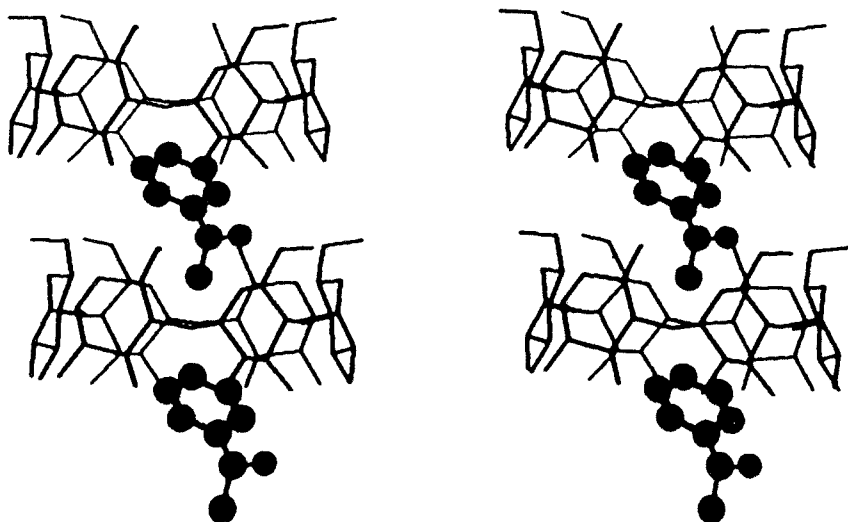
Fig. 1. Stereoscopic representation of the crystallographic structure of the α-cyclodextrin/phenyl-ethanol complex.

For simplification, both backbones (non-hydrogen atoms) were kept rigid in this first step because the flexibilities of the two moieties are unimportant with regard to the global geometry of the system. The hydroxylic hydrogens were omitted in order to suppress minor flexibilities at this stage. Thus, the only internal parameters taken into account were those of the guest positioning.

The relative position and orientation of the guest molecule are described by six parameters (p1 to p6) internal to the CD complex, involving a distance (p1), an angle (p2) and four dihedrals (p3 to p6). Virtual atoms (e.g. molecule centroids) and virtual bonds such as normals to the least squares planes of the molecules were used for the definition of these parameters. This crankshaft technique [34] was preferred to more classical docking procedures that use three parameters for translation and three others for rotation [35] or Euler angles [36] since dihedrals are easier to manage and to visualize with SYBYL. But the clustering scheme described below remains valid for any kind of parameters.

A schematic representation of the two molecules and the positioning parameters is shown in Figure 2. In this construction, G and G′ are the centroids of the non-hydrogen atoms of the αCD and of the phenyl moiety of the PE, respectively. GI is the normal to the least squares plane calculated from the skeleton of the αCD. Two virtual bonds, G′H′ and H′I′, were created with very short lengths (0.01 nm) to induce an almost perfect rotation (without any significant precession movement) centered on G′. Finally the virtual atom I′ was bonded to the C1 atom of PE.

The six parameters for the systematic scanning are:

p1: distance GG′,

p2: angle IGG′,

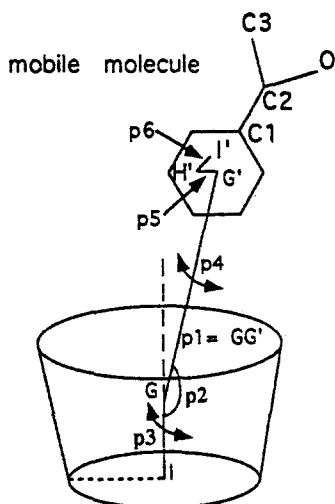p3: dihedral (rotatable bond in the SYBYL terminology) around IG,

Fig. 2. Schematic description of the six parameters (p1, p2, p3, p4, p5, p6) used for systematic scanning of the cyclodextrin cavity by the guest molecule.

p4: dihedral around GG',

p5: dihedral around G'H',

p6: dihedral aroung H'I'.

The isomerism of PE is present in the crystal structure with a statistical disorder. Only the (−) form was selected for calculations, since the orientation of the hydroxyl group plays a minor role in the docking aspect. Complexation can occur on either the primary or the secondary faces (those with the primary alcohol groups and smaller diameter and those with the O2 and O3 groups and larger diameter, respectively). Obviously, both types are observed by crystallography (Figure 1) or more recently by NMR spectroscopy [37]. The calculations were divided into two parts (Figure 3): complexation on the secondary side (A docking type) and the primary side (B docking type). For technical reasons explained below, the first type was divided in two cases (A' and A" docking sub-types) with, respectively, the phenyl ring or the ethanol part of the guest pointing towards the cavity.

Scanning was performed with the SEARCH routine [38, 39] of SYBYL, which can only manage dihedrals. Therefore, each pair of p1 (distance) and p2 (angle) parameters was treated separately with complete torsion angle scan. The results of all these calculations were merged with a procedure external to SYBYL. Preliminary calculations determined the ranges and the increments used for the six parameters. All the chosen values are summarized for the two types of docking (A and B) in Table I. As can be seen, the differentiation between A' and A" sub-types is only due to p1 values. The limits of the parameter p1 are the most difficult to determine; the lower limit was obtained when the SEARCH calculations yielded no more solutions due to steric conflicts by further decreasing the p1 value. The upper limit was taken as the distance above which the total energy of the system is
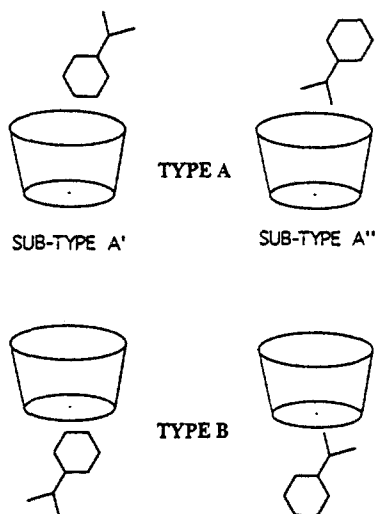
Fig. 3. Schematic representation of the A', A" docking sub-types and B docking type.

Table I. Variation ranges used for the six positioning parameters in the SEARCH calculations for A', A" docking sub-types and B docking type (parameter increments are in brackets).

	sub-type A'	sub-type A"	type B
p1 (nm)	0.29 (0.1) 0.38	0.4 (0.1) 0.46	0.38 (0.2) 0.58
p2 (°)	150 (5) 180	150 (5) 180	0 (5) 30
p3 (°)	185 (5) 265	185 (5) 265	0 (5) 120
p4 (°)	0 (20) 340	0 (20) 340	0 (20) 340
p5 (°)	0 (20) 340	0 (20) 340	0 (20) 340
p6 (°)	0 (20) 340	0 (20) 340	0 (20) 340

constant. The ranges of p2 and p3 were reduced for obvious steric conflicts and sixth order pseudo-symmetry of the rigid αCD. The parameters p4, p5 and p6 were scanned over 360°. All the redundant solutions due to scanning and symmetries were eliminated *a posteriori*.

The most improbable structures were excluded with standard options of SYBYL using both steric (van der Waals term) and energetic criteria (energy cut off). The van der Waals contribution is very coarse because a rigid model was used. Therefore the van der Waals radii were reduced by 0.85 to allow for possible local rearrangements that would resolve small steric conflicts. Structures with energies 20.9 kJ mol^{-1} (5 kcal mol^{-1}) above the lowest energy conformation of each docking type (A and B) were cut off.

2.2. FAMILIES OF LOW-ENERGY STRUCTURES

A general problem in docking procedures is the extraction of a few representative host–guest interaction types from the large number of results obtained by systematic

scanning of the relative guest positions. The program PROXIM was written for the comparison of low-energy structures and their classification into families. It is an extension of a clustering procedure successfully applied in a previous systematic conformational search [40] and describes the resulting families in terms of energetic, geometrical or topographical criteria. From this classification, one or several structures can be selected as characteristic of each family for a more complete molecular modelling analysis, leading to possible structures of the corresponding complexes.

This program uses a very simple definition of proximity based on internal parameters. The proximity of two solutions is defined as the sum of relative differences between normalized values of the p1–p6 parameters. For each parameter the increments for the systematic scanning calculations (Table I) were taken as unit values (normalization) for the proximity calculation. With this definition, both proximity values and proximity criterion are integers. Then, two solutions are classified in the same family if the calculated proximity is smaller than, or equal to, the proximity criterion. If, for example, p1 through p5 are equal for two solutions, but p6 is different by only one increment, the two structures will be in the same family if the proximity criterion is one. If the criterion is two, then p2 and p6 could each differ by one and the two solutions would be in the same family. Alternatively, with a criterion of two, p4 could differ by two increments. For the classification of a great number of solutions, this sum is much faster to compute than root mean square (RMS) values, especially when the number of parameters is important. A schematic representation of a clustering based on two parameters and a proximity criterion equal to 1 is shown in Figure 4.
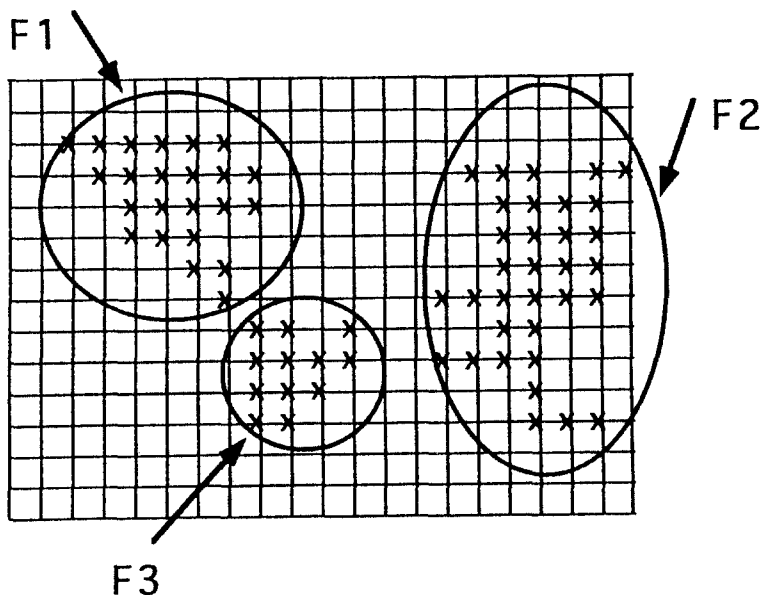

Fig. 4. Illustration of the clustering method as calculated by the PROXIM algorithm in a two dimensional space and with a proximity distance set to 1.

PROXIM is similar to the FAMILY option of the SYBYL package but more versatile since it is possible to adjust the proximity criterion value. Moreover, PROXIM supplies additional information such as comparison between the lowest energy and the average solutions, analysis of mean and maximal deviations for each parameter, topographical information such as general shape, convexity or compactness, analysis of the envelopes of the three inertia moments and proximity between the average solutions of the different families are also calculated. The lowest energy solution of a family is chosen as representative of a docking result if it is close to the average solution of the family, and if the set of solutions is sufficiently convex and compact. Otherwise, several representative solutions must be selected, using energetic, geometrical criteria given by PROXIM, in order to define more compact and coherent sub-families. For the example shown in Figure 4, the use of a proximity criterion equal to 2 would lead to the gathering of F1 and F3 in the same family, a criterion of 3 would combine F1, F2 and F3 in the same family. Such families would, however, be less compact and less convex. After clustering, the docking results characteristic of families were displayed to identify redundancies due to the scanning conditions and the symmetries of host and guest molecules.

Three 'synthetic' parameters s1, s2 and s3 were used to describe the docking results with the smallest possible number of parameters (Figure 5). The first parameter (s1) is the angle between the αCD and the phenyl ring least squares planes and measures the tilt of the guest with regard to the host (Figure 5a). The second parameter (s2) gives the orientation of the alcohol group of the PE molecule (Figure 5b). The third parameter (s3) gives the orientation of the carbon atom C_3 of the ethanol group. As shown in Figure 5c, this atom has two positions (above and below) with regard to the oxygen atom. However (s3) has no meaning for the structures corresponding to s1 about 90° for which the two atoms O and C_3 have similar positions.
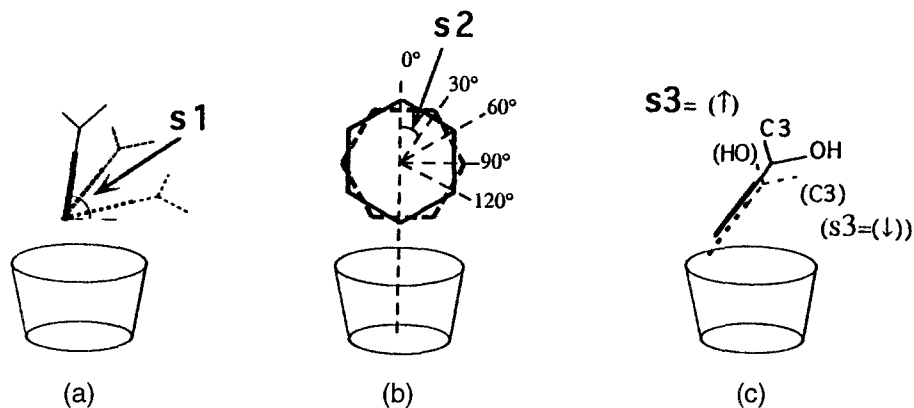

(a)   (b)   (c)

Fig. 5. Schematic representations of the three parameters s1, s2, s3 used for the classification of families. 5a; Tilting of the phenyl-ethanol molecule according to the different values determined for the s1 parameter. 5b; Relative orientations of the C1–C2 bond of phenyl-ethanol with regard to values of the s2 parameter. 5c; Relative position ('above' and 'below') of 'C3' versus 'O' atom as described by the parameter s3, this parameter is irrelevant when s1 = 90°.

2.3. REFINEMENT OF THE BEST DOCKING RESULTS

The lowest energy structures representing the different families or sub-families (docking results) were refined in two successive steps. During the first molecular mechanics minimization with the MAXIMIN2 [41, 42] module of SYBYL, all the internal parameters were relaxed to evaluate the rearrangement of the skeletons of both molecules (e.g. the deformation of αCD symmetry). After the addition of the missing hydroxylic hydrogens, a second minimization was performed with partial charges calculated with the Pullman method [43]. Similar calculations using the Gasteiger–Hückel approximation [44] yielded the same results. This ultimate minimization was necessary to take into account the side flexibility of the primary hydroxyl groups and the electrostatic interactions between host and guest molecules. In order to check the validity of the two-step refinements, direct minimizations from the SEARCH results were also carried out.

3. Results and Discussion

3.1. SYSTEMATIC SCANNING

The number of theoretical relative positions of the guest molecule explored by the complete scanning of the six parameters, and the number of selected positions after exclusion of the most improbable results are given in Table II for the docking types. The percentages of selected structures obtained were of the same order of magnitude except for the A″ sub type. Nevertheless, the total number of solutions (28944–0.13%) was still too large to be easily interpreted.

3.2. FAMILIES OF LOW ENERGY STRUCTURES

PROXIM was used separately for each docking type. After some preliminary trials, the proximity criterion was set to 1. This criterion gave about 100 satisfyingly compact families, each consisting of several hundred members. In all cases, the average and the lowest energy structures were similar in terms of location in the multidimensional space (p1, p2, . . . ,p6) as well as in terms of energy. The validity of the partitioning into the different families was examined using molecular graphics representation of families.

Due to different redundancies, identical families were found and discarded. The first kind of redundancy was directly due to the over large scanning range used for

Table II. Total number of explored positions by systematic scanning of the parameters p1 to p6 and selected solutions for A′, A″ sub-types and B type.

	sub-type A′	sub-type A″	type B
Calculated Solutions	6 940 080	4 858 056	11 226 600
Selected Solutions	12321 (0.18%)	790 (0.02%)	15833 (0.14%)

the parameters p4, p5 and p6. Two families (f and f') are identical when their parameters are related by the following relationships:

$$p4(f') = p4(f) - 180°$$

$$p5(f') = p5(f) \qquad \text{(modulo 360°)}$$

$$p6(f') = p6(f) + 180°$$

The second redundancy corresponded to the degeneration of the parameter p2: when p2 = 180° (or p2 = 0°), two families having the same sum (p3 + p4) are identical. The last redundancy was specific to the host–guest shape and involved the sixth order symmetry of the αCD in its rigid form. If p2 = 180° (or p2 = 0°), the solutions are exactly centered and the five other equivalent solutions are available, even with a partial scanning of the parameter p3. If p2 ≠ 180°, any solution must have, at most, one equivalent since the variation range of p3 is larger than 60°.

After discarding the redundant families, there remained 30 families of docking results (10 and 20 for the A and B docking types, respectively). Since each family can be represented by a unique docking result, PROXIM led to a severe reduction from the 28944 selected solutions to 30 docking results (0.1%).

The classification of these families using the parameters s1, s2 and s3, easily related to geometrical features of the host–guest associations, is summarized in Table III. In this table, the three A″ sub-type structures (A3, A8 and A10) are characterized by s2 values of 150 or 180°. The numbering of the families inside each docking type was chosen according to an arbitrary hierarchy of the parameters (s1 > s2 > s3). The 30 families could be classified into three groups based on the s1 values (Figure 5a). The first group of 8 structures (denoted as (1) in s1 column of Table III) corresponds to a tilt of about 90° (80° < s1 < 90°), the phenyl ring being nearly perpendicular to the axis of the host cavity. The 14 structures of the second group (2) have an intermediate tilt value (40° < s1 < 60°) while for those (8 structures) of the third group (3) the phenyl ring is almost parallel to the αCD plane (10° < s1 < 20').

The observed values of s2 in the final structures are respectively 0°, 60°, 120°, 180° and 30°, 90°, 150°. The corresponding orientations of the phenyl ring are shown in Figure 5 for the A docking type. As for the αCD symmetry, the capability of PROXIM to segregate families only separated by the rotation of the alcohol group (along the normal to the phenyl ring), demonstrates the robustness of the clustering method.

Both relative orientations were found for the C_3 and O atoms (Figure 5c) as shown by the up and down arrows in the s3 column. No arrow is shown for the structures corresponding to s1 about 90° because the two atoms O and C_3 have similar positions.

All the combinations of s1, s2 and s3 values were not found on account of obvious steric effects. For instance, the combination s1 ∼ 90° (1), s2 = 90° for the A docking type would correspond to impossible interactions between some atoms of the αCD and that of the alcohol group of the PE. The steric conflicts were especially important for docking A″ sub-type where only 3 structures (A3, A8, A10) were found, each with different tilt values of s1.

Table III. Description of the selected families in terms of synthetic parameters and relative energies. The symbol (*) indicates the members of the A″ sub-type. The A4 and B6 docking solutions (annotated "crystal") refer to PE sandwiched between two cyclodextrin hosts.

Parameters. s1: (1), (2) and (3) correspond to three categories represented in Figure 5a. s2: orientation of the C1–C2 bond (Figure 5b). s3: the arrows illustrate the relative positions ("up" and "down") of the C3 atom (Figure 5c).

Engergies. I: rigid entities (SEARCH); II: relaxation of skeletons; III: final refinements with partial charges and hydroxylic hydrogens.

Conf. No.	Parameters			Energies (kJ mol^{-1}) or {kcal mol^{-1}}			
	S1	S2 (°)	S3	I	II	III	
A1	(1)	30		15.9{3.8}	15.5{3.7}	17.6{4.2}	
A2	(1)	60		18.4{4.4}	16.3{3.9}	19.7{4.7}	
A3*	(1)	180		11.7{2.8}	0.0{0.0}	10.0{2.4}	
A4	(2)	0	↑	18.8{4.5}	10.5{2.5}	8.4{2.0}	crystal
A5	(2)	30	↑	15.1{3.6}	8.8{2.1}	18.8{4.5}	
A6	(2)	30	↓	14.2{3.4}	13.4{3.2}	19.7{4.7}	
A7	(2)	60	↑				
A8*	(2)	180	↓	0.0{0.0}	1.3{0.3}	0.4{0.1}	
A9	(3)	30	↑	15.9{3.8}	17.2{4.1}	19.3{4.6}	
A10*	(3)	150	↑	14.7{3.5}	20.1{4.8}	13.0{3.1}	
B1	(1)	0		18.4{4.4}	23.9{5.7}	25.5{6.1}	
B2	(1)	30		20.5{4.9}	28.9{6.9}	24.7{5.9}	
B3	(1)	90		23.0{5.5}	30.6{7.3}	27.9{6.6}	
B4	(1)	150		24.3{5.8}	28.5{6.8}	26.0{6.2}	
B5	(1)	180		23.9{5.7}	28.5{6.8}	25.5{6.1}	
B6	(2)	0	↑	8.8{2.1}	14.7{3.5}	10.0{2.4}	crystal
B7	(2)	0	↓	14.7{3.5}	11.3{2.7}	21.3{5.1}	
B8	(2)	30	↑	23.9{5.7}	15.9{3.8}	9.6{2.3}	
B9	(2)	30	↓	21.3{5.1}	31.8{7.6}	19.7{4.7}	
B10	(2)	90	↓	23.9{5.7}	26.4{6.3}	32.6{7.8}	
B11	(2)	120	↓	25.1{6.0}	11.7{2.8}	0.0{0.0}	
B12	(2)	150	↑	19.7{4.7}	7.5{1.8}	20.1{4.8}	
B13	(2)	180	↑	24.3{5.8}	20.1{4.8}	18.0{4.3}	
B14	(2)	180	↓	24.7{5.9}	27.6{6.6}	26.4{6.3}	
B15	(3)	30	↑	22.6{5.4}	1.7{0.4}	14.7{3.5}	
B16	(3)	30	↓	22.6{5.4}	20.9{5.0}	17.6{4.2}	
B17	(3)	60	↑	24.3{5.8}	7.5{1.8}	8.4{2.0}	
B18	(3)	60	↓	24.3{5.8}	15.1{3.6}	24.7{5.9}	
B19	(3)	150	↓	18.0{4.3}	25.5{6.1}	28.5{6.8}	
B20	(3)	180	↓	24.3{5.8}	32.6{7.8}	33.5{8.0}	

3.3. REFINEMENTS OF THE BEST DOCKING RESULTS

In order to validate the quality of the SYBYL force field, Harata's PE-αCD X-ray structure was refined with its current force field. Superimposition of the two structures yields a RMS deviation for all non-hydrogen atoms of 0.0204 nm. This value was considered as acceptable, taken into account that no crystal constraint was imposed. The 30 docking results were optimized with molecular mechanics. The corresponding relative energies are given in Table III (right part), with an arbitrary

zero value taken as the lowest energy structure calculated for each refinement step. The initial energies of the complex (skeletons of αCD and PE) as obtained after the SEARCH procedure are in column I. At this stage, both moieties were considered as rigid. The lowest energy structure is the one found for A8. A relative energy higher than 25.1 kJ mol^{-1} (6 kcal mol^{-1}) (e.g. B11) is not incompatible with the used cut-off of 20.9 kJ mol^{-1} (5 kcal mol^{-1}) since the scannings were performed separately for the A and B types.

The energies obtained after the first minimization are reported in column II. For a given docking result, the energy evolution from I to II measures the relaxation of the skeletons which removed some local steric conflicts. At this stage, the lowest energy structure was again A8. There is no important alteration of the αCD symmetry.

In column III are the best values obtained either after the second minimization procedure or a direct minimization from the SEARCH results. Both of these optimizations take into account the partial charges and the contribution of the hydroxyl groups to the stabilization of the structure. For each docking type or sub-type, at least one very probable structure was found (A4, A8 and B11 are presented in Figure 6). Others may be considered (A3, B6, B8 and B17) for A and B types. The refined structures were close to the initial ones and no fusion between the docking results occurred during these refinements, except for A7 and A2. In this case, the refinements of A7 provoked a variation of the tilt value s1 so that the final conformation was similar to that of A2. This means that the clustering algorithm yielded fairly well segregated solutions and that the refinements only diminished some local steric conflicts. Thus, the characterization of the minimized structures (representative of the docking results) with the synthetic parameters s1, s2 and s3 remains valid.

The general crystal packing conditions require that s1, s2 and s3 are identical in A and B docking types, respectively. The A4 and B6 results correspond to the two types of host/guest association present in the crystal structure (Figure 1) for which the relative position of PE is measured by s1 = 45°, s2 = 2° and s3 $\equiv$ up. Although they are not the lowest energy structures, they were among the low energy solutions (A4: 8.4 kJ mol^{-1} (2.0 kcal mol^{-1}) and B6: 10.0 kJ mol^{-1} (2.4 kcal mol^{-1})). However, their combination was the best one in terms of total energy (18.4 kJ mol^{-1} (4.4 kcal mol^{-1})) when comparing structures of the A and B docking types. The four other possible combinations with respect to the crystal packing conditions (complexation of both primary and secondary faces) were (A8, B14), (A5, B8), (A9, B15) and (A3, B5). Those pairs of structures have higher total energies (26.8 kJ mol^{-1} (6.4 kcal mol^{-1}), 28.5 kJ mol^{-1} (6.8 kcal mol^{-1}), 33.9 kJ mol^{-1} (8.1 kcal mol^{-1}) and 33.6 kJ mol^{-1} (8.5 kcal mol^{-1}), respectively).

The extent of relaxation of the internal parameters during the final refinement is revealed by the deformations of the αCD symmetry. Among different characterizations of this deformation, van Helden *et al.* [45] have suggested the ratio $Dx = d_{max}/d_{min}$ of facing glycosidic oxygen distances. In order to differentiate the deformations of the secondary and the primary faces, we have measured the distances d_{C2} and d_{C5} along the diagonals C-2$_{(n)}$...C-2$_{(n+3)}$ (secondary face) and C-5$_{(n)}$...C-5$_{(n+3)}$ (primary face) of the αCD. Table IV summarizes the mean maximum and minimum values of d_{C2} and d_{C5} in each docking type. As expected,
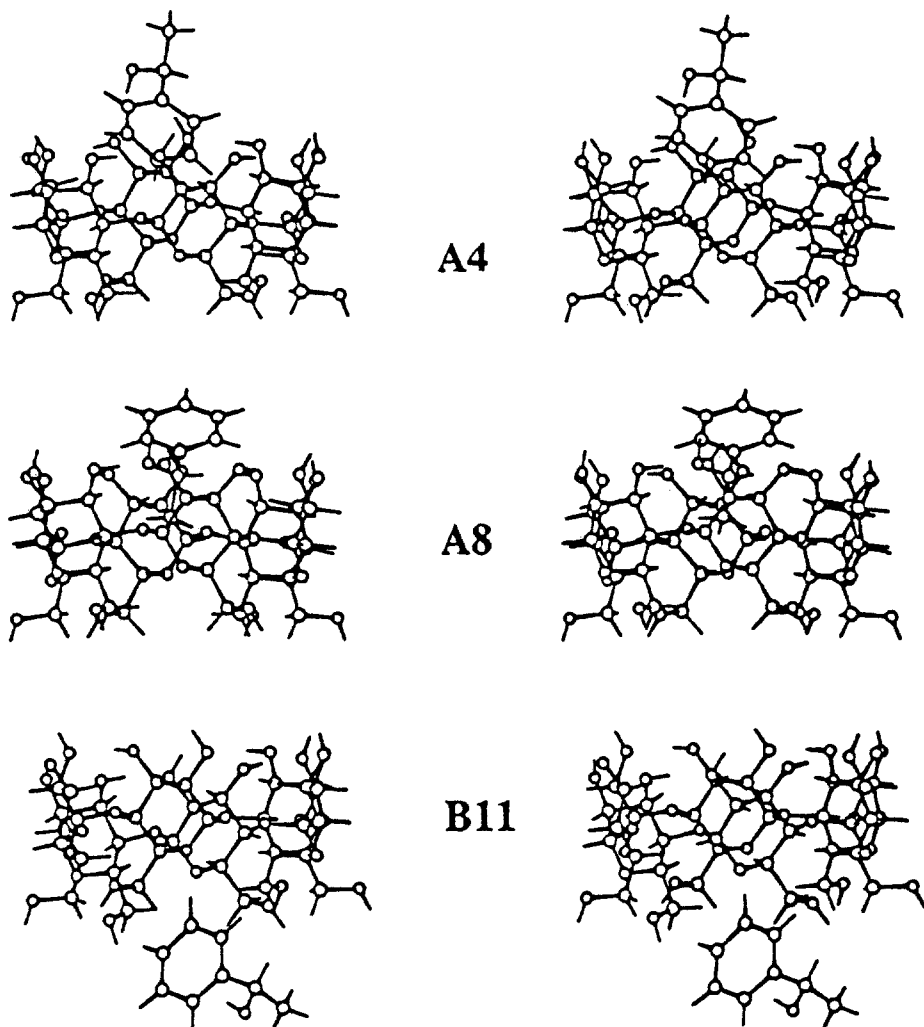
Fig. 6. Stereoscopic representations of the lowest energy conformations obtained in A', A'' docking sub-types and B docking type.

Table IV. Deformation of the cyclodextrin ring due to complexation. In columns 1 to 4 are given the mean maximum and minimum values (nanometers) for the diagonals $C2_n - C2_{n+3}$ and $C5_n - C5_{n+3}$. D_{C2} and D_{C5} are the corresponding d_{max}/d_{min} ratios.

	d_{C2max}	d_{C2min}	d_{C5max}	d_{C5min}	D_{C2}	D_{C5}
type A						
(A')	1.140	1.061	0.870	0.859	1.07	1.01
(A'')	1.136	1.075	0.870	0.862	1.06	1.01
type B	1.134	1.084	0.878	0.860	1.05	1.02

the differences between $d_{C2_{max}}$ and $d_{C2_{min}}$ are higher for the A′ sub-type (inclusion of the phenyl ring, secondary face) than for the A″ sub-type (inclusion of the ethanol part, secondary face) and for the B type (inclusion in the primary face). These values are respectively 0.079 nm, 0.061 nm and 0.050 nm. The variations observed for d_{C5} indicate that the primary faces are less perturbed in the A docking type than in the B docking type. The corresponding values are respectively 0.010 nm, 0.008 nm and 0.018 nm. These last results show that the proposed strategy is able to take into account some eventual deformation even if the procedure works with rigid entities in the initial stage.

4. Conclusion

The described methodology is suitable for other docking problems by adjustment of the six positioning parameters following specific features of both host active part and guest molecule. With a simple definition of proximity, PROXIM is available for clustering of solutions in any type of systematic search of conformations or interactions, so long as suitable normalized units are chosen for the definition of the proximity criterion. However, since this criterion can take different values, the result must be confirmed by the topographic information given by the program and by graphical representations.

This methodology was applied to the αCD/PE complex for which a crystal structure is known. The pseudo-symmetries of both molecules were an excellent test for the robustness of the PROXIM program and a limited number of structures were obtained and visualized as distinct docking cases. The two molecules were initially considered as rigid entities, then allowed to relax in the final refinements. This relaxation induced deformations of the αCD symmetry, comparable in size to experimentally determined structures in the literature.

Thus, this methodology seems to be powerful for detection and analysis of possible stable complexes. It is presently tested with more than six parameters (six positioning parameters plus some other intramolecular internal parameters). Thus, this approach could be generalized taking into account some flexibility of the two entities at the scanning stage.

References

1. C. A. Venanzi, P. M. Canzius, Z. Zhang and J. D. Bunce: *J. Comput. Chem.* **10**, 1038 (1989).
2. D. W. Griffith and M. L. Bender: *Adv. Catal.* **23**, 209 (1973)
3. R. Breslow, M. F. Czarniecki, J. Emert and H. Hamaguchi: *J. Am. Chem. Soc.* **102**, 762 (1980).
4. Y. Inoue, T. Okuda and Y. Miyata: *Carbohydr. Res.* **192**, 17 (1989).
5. M. L. Bender and M. Komiyama: *Cyclodextrin Chemistry*, Springer Verlag, Berlin (1978).
6. J. Szejtli: *Cyclodextrins and Their Inclusion Complexes*, Akademiai Kiado, Budapest (1982).
7. R. J. Clarke, J. H. Coates and S. F. Lincoln: *Adv. Carbohydr. Chem.* **46**, 205 (1988).
8. M. Kitagawa, H. Hoshi, M. Sakurai, Y. Inoue and R. Chujo: *Bull. Chem. Soc. Jpn.* **61**, 4225 (1988).
9. F. Cramer and W. Kampe: *J. Am. Chem. Soc.* **87**, 1115 (1965).
10. K. Harata: *Bull. Chem. Soc. Jpn.* **49**, 2066 (1976).
11. Y. Matsui: *Bull. Chem. Soc. Jpn.* **55**, 1246 (1982).
12. I. Tabushi, Y. Kiyosuke, T. Sugimotro, R. Yamamura: *J. Am. Chem. Soc.* **100**, 916 (1978).
13. R. J. Bergeon and R. Rowan: *Bioorg. Chem.* **5**, 425 (1976).
14. W. Saenger: *Magn. Reson. Chem.* **24**, 835 (1986).

15. C. Betzel, W. Saenger, B. E. Hingerty and G. M. Brown: *J. Am. Chem. Soc.* **106**, 7545 (1984).
16. J. A. Ripmeester: *J. Incl. Phenom.* **4**, 129 (1986).
17. M. G. Usha and R. J. Wittebort: *J. Am. Chem. Soc.* **114**, 1541 (1992).
18. M. Komiyama and H. Hirai: *Bull. Chem. Soc. Jpn.* **54**, 828 (1981).
19. A. W. Coleman, G. Tsoucaris, H. Parrot, H. Galons, M. Miocque, B. Perly, N. Keller and P. Charpin: *J. Chromatogr.* **450**, 175 (1988).
20. G. Le Bas, N. Rysanek, F. Royer and G. Tsoucaris: *Mol. Cryst. Liq. Cryst. Inc. Nonlin. Opt.* **161**, 363 (1988).
21. I. Lukovits: *J. Mol. Struct. (Theochem)* **170**, 249 (1988).
22. M. J. Sherrod: *Carbohydr. Res.* **192**, 17 (1989).
23. E. N. Arnold, T. S. Lillie and T. E. Beesley: *J. Liq. Chromatogr.* **12**, 337 (1989).
24. M. Sakurai, M. Kiatagawa, H. Hoshi, Y. Inoue and R. Chujo: *Carbohydr. Res.* **198**, 181 (1990).
25. J. E. Koehler, W. Saenger and W. F. van Gunsteren: *J. Biochem. Struct. Dynamics* **6**, 181 (1988).
26. J. E. Koehler, W. Saenger and W. F. van Gunsteren: *Eur. Biophys. J.* **15**, 197 (1987).
27. M. Prabhakaran: *Biochem. Biophys. Res. Commun.* **178**, 192 (1991).
28. S. H. Nilar: *J. Comput. Chem.* **12**, 1008 (1991).
29. K. B. Lipkowitz, D. A. Demeter, R. Zegarra, R. Larter and T. Darden: *J. Am. Chem. Soc.* **110**, 3446 (1988).
30. M. G. Still and L. B. Rogers: *J. Comput. Chem.* **11**, 242 (1990).
31. M. Blanco: *J. Comput Chem.* **12**, 237 (1991).
32. K. Harata: *Bull. Chem. Soc. Jpn.* **55**, 1367 (1982).
33. *SYBYL: Molecular Modelling System*, Tripos Associates, St Louis, MO.
34. A. Imberty, K. D. Hardman, J. P. Carver and S. Pérez: *Glycobiology* **1**, 631 (1991).
35. M. G. Rossmann and D. M. Blow: *Acta Crystallogr.* **15**, 24 (1962).
36. J. Janin and S. Wodak: *Biopolymers* **24**, 509 (1985).
37. F. Djedaïni: '*Etude par résonance magnétique nucléaire des phénomènes d'inclusion et d'adaptation moléculaire dans les cyclodextrines naturelles et des dérivés synthétiques*', Thesis, Paris (1991).
38. I. Motoc, R. A. Dammkoehler, D. Mayer and J. Labanowski: *Quant. Struc. Act. Relat.* **5**, 138 (1986).
39. I. Motoc, R. A. Dammkoehler and G. O. Marshall in *Mathematical and Computational Concepts in Chemistry* (Ed. N. Trinajestic), Ellis Horwood Ltd., p. 222 (1986).
40. A. Buléon and V. Tran: *Int. J. Biol. Macromol.* **12**, 345 (1990).
41. I. Motoc, R. A. Dammkoehler, D. Mayer and J. Labanowski: *Quant. Struc. Act. Relat.* **5**, 99 (1986).
42. J. G. Vinter, A. Davis and M. R. Saunders: *J. Comput. Aid. Mol. Design* **1**, 31 (1987).
43. H. Berthod and A. Pullman: *J. Chem. Phys.* **62**, 942 (1965).
44. J. Gasteiger and M. Marsili: *Tetrahedron* **36**, 3219 (1980).
45. S. P. van Helden, M. J. van Drooge, A. J. Claessens, A. C. Jansen and L. H. M. Janssen: *Carbohydr. Res.* **215**, 251 (1991).