ELSEVIER

Carbohydrate Research 276 (1995) 267–287

CARBOHYDRATE RESEARCH

A self-included cyclomaltoheptaose derivative studied by NMR spectroscopy and molecular modelling

Patrick Berthault [a,*], Delphine Duchesne [a], Hervé Desvaux [a], Bernard Gilquin [b]

[a] *CEA, Centre d'Etudes de Saclay, DRECAM / Service de Chimie Moléculaire, F-91191 Gif-sur-Yvette, France*

[b] *CEA, Centre d'Etudes de Saclay, Département d'Ingéniérie et d'Etude des Protéines, F-91191 Gif-sur-Yvette, France*

Received 1 February 1995; accepted 4 May 1995

Abstract

The 3D structure of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose, a self-complexing β-cyclodextrin derivative, was determined by NMR and molecular modelling. The aminoacyl side-chain is included in the cavity and induces chemical-shift variations in the CD proton signals, allowing their complete assignment. Dipolar interactions between protons of the tyrosine ring and internal protons of the cyclodextrin were used to obtain distance constraints. Then 42 structures were calculated from 32 distance constraints — 21 shorter than 4 Å involve the host–guest interactions — using a simulated annealing procedure. Starting from one of the resulting structures, a 250-ps molecular dynamics simulation was carried out in a waterbox without constraint. The simulation data are in agreement with NMR data such as nOe and ring-current effects. The cyclodextrin part takes an elliptical shape, which tightly fits the aromatic moiety. As a consequence, the respective motion of the host and the guest moieties have the same amplitude and time scale: the self-inclusion complex shows only little flexibility.

Keywords: Cyclodextrin; Intramolecular complex; Self-inclusion; ^{1}H and ^{13}C NMR; Molecular dynamics; Ring currents; nOe restraints; Simulation in waterbox

Abbreviations: SA: simulated annealing; rms: root mean square; MD: molecular dynamics; CD: cyclodextrin; nOe: nuclear Overhauser effect.

* Corresponding author.

SSDI 0008-6215(95)00193-X

1. Introduction

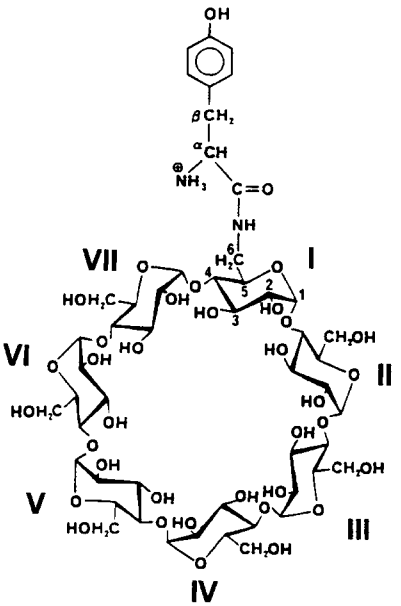
Cyclodextrins are natural cyclic oligosaccharides constituted mainly by six (cyclomaltohexaose, α-cyclodextrin, α-CD), seven (cyclomaltoheptaose, β-cyclodextrin, β-CD) and eight (cyclomaltooctaose, γ-cyclodextrin, γ-CD) D-glucose units linked in α-(1 $\rightarrow$ 4). Their toroidal shape is schematically delimited by the primary hydroxyls (OH-6) for the narrow side, and by the secondary hydroxyls (OH-2, OH-3) for the wide rim. Their ability to accommodate hydrophobic guest molecules into their cavity allows one to solubilize the resulting inclusion complexes in water, and offers a wide range of applications [1–3]. Among them, chiral recognition seems to represent the most interesting property for chemistry, and vectorization of hydrophobic drugs the most striking challenge for biological and pharmaceutical purposes. However, in some cases, it was pointed out that the native cyclodextrin reveals a poor ability for optical isomer separation [4]. For this reason, as well as to decrease the toxicity level, to increase the host–guest association constant and to improve drug administration properties, modifications of the CDs by selective substitutions on the primary or secondary hydroxyls are nowadays under development in many laboratories [4–10]. For example, grafting on the primary side a peptidic fragment—designed to recognize a particular cellular receptor—is currently in progress for site-directed drug delivery purposes [11].

The choice of the first residue attached to C-6 is essential for obtaining a self-inclusion complex. When tyrosine or phenylalanine is grafted on to β-cyclodextrin, the aromatic ring is trapped into the hydrophobic cavity whatever the enantiomeric configuration of the amino acid [12–17]. According to the nature of the side chain and the chirality of the residue, the aromatic moiety can adopt different positions, and the inclusion of one enantiomer is favoured in comparison with the other one. Therefore these compounds showing improved ability of chiral recognition are excellent models for understanding the inclusion process, and for providing insight into the deformations of the cavity during the molecular adaptation.

In this paper, the 3D structure as well as the conformational and the dynamic properties of the self-inclusion complex 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose (**1**) were determined using dipolar relaxation data and molecular dynamics calculations. This compound was previously synthesized and studied by NMR by Takahashi [15,16] and Perly and co-workers [17], but only a rough ^{1}H spectral assignment and no structural determination were performed.

2. Experimental

NMR spectroscopy.—All NMR experiments were performed at 278 K on a standard Bruker AMX500 spectrometer equipped with a cooling unit for temperature regulation. The sample was constituted by 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose (25 mg) dissolved in $^{2}H_2O$, giving a 48 mM solution. Chemical shifts were referenced relative to 4,4-dimethyl-4-silapentanoic acid. For all the semi-soft two-dimensional experiments, 128 increments of 2048 points were collected in TPPI mode [18]. After zero-filling to



1

1024 points in the second dimension F1, the time domains were treated by a sinebell window shifted by $\pi/4$ in both dimensions and Fourier transform. The inverse heteronuclear correlations were constituted of 380 FIDs of 2048 points (96 scans), zero-filled to 1024 points in F1 and Fourier transform after filtering by cosinebell functions in both dimensions. The pseudo-3D ROESY-TOCSY experiment, designed to confirm the previous assignment, was constituted by 256 FIDs of 2048 points zero-filled to 1024 in F1 before Fourier transformation. The mixing times for the ROESY and TOCSY parts were 200 ms and 20.7 ms, respectively.

Seven 2D NOESY experiments in D_2O were performed at different mixing times — from 50 to 400 ms — in order to estimate as accurately as possible interproton distances from cross-relaxation measurements through the initial build-up rate approximation [19]. The acquired matrices were made up of 256 FIDs of 2048 points and were zero-filled to 1024 points in F1 before Fourier transformation. No filtering window was applied. A 2D NOESY experiment with a mixing time of 300 ms was performed in light water by the Jump–Return method [20]. The acquisition and processing parameters were the same as for the other NOESY. Integration of the cross-peak volumes in NOESY maps was performed in Bruker/UXNMR software with rescaling of the measured value relatively to the diagonal peak. Thirty-two nOe's between the amino acyl linkage protons and protons of cyclodextrin were recorded and transformed into distance data, after fitting the build-up curves by a second order polynom. Most of them, 21, are shorter than 4 Å. The average H-1–H-4 dipolar interaction was chosen as a reference, as it presented the advantage of showing no distortion inherent to scalar coupling and no important fluctuation between different interglycosidic linkages.

Measurement of the effective local correlation times by off-resonance irradiation was performed with the adiabatic version of the sequence [21,22]. Thirty-three experiments were carried out with different values of the angle between the static and the effective rf magnetic field regularly spaced between 4 and 57°. The irradiation delay was 1.3 s and the rf field strength 2.3 G.

Measurement of each H-1–C-4 coupling constant was performed using the method of Freeman et al. [23], as it can be related to the interglycosidic torsion angle by a semi-empirical Karplus-type relationship [24]. However, no significant variation between different unit linkages was found out of the error range, and therefore no relevant structural information was deduced from these measures for molecular modelling.

Molecular modelling.—The three-dimensional structure of this self inclusion complex was determined by a simulated annealing procedure using nOe data as restraints, then a free molecular dynamics in a waterbox starting from one of the obtained structures was performed.

(i) Simulated annealing using NMR data. The calculations were performed in vacuum with the software X-PLOR 3.1 [25]. The target function was the same as that used by Nilges and co-workers [26]. The non-bonded interactions were taken into account by a repulsive term in place of the van der Waals and electrostatic terms. In a first stage, the structures were generated by ab-initio simulated annealing (SA) procedure. Starting from a template conformation (randomly generated, but exhibiting non-aberrant geometrical values), a first 6-ps dynamics with a timestep of 2 fs at 1000 K was followed by a 6-ps slow cooling to 100 K. The force constant on the nOe restraints was increased progressively in the high temperature dynamics. A "soft" shape of the nOe restraint function defined by a low-slope asymptote far from the minimal value allowed a progressive fit of the constraints. A small weight of the van der Waals repulsive term was used at high temperature in order to enable a large conformational sampling. Fifty structures were generated by repeating this protocol. In a second stage, the resulting structures were refined by the following procedure: 2 ps simulated annealing from 2000 to 100 K (timestep 3 fs), with a square shaped nOe potential function and a progressive increase of the van der Waals repulsive term to its usual value. The force constant on the nOe term was set to 50 kcal/Å^2. The final conformations were minimized, analyzed and selected using acceptance criteria on the constraints and covalent geometry terms.

Hδ and Hδ' NMR signals being equivalent, as well as Hϵ and Hϵ', the pseudo-atom facility of X-PLOR was used. Averaging in R^{-6} for interproton distance restraints was chosen. For ambiguous dipolar cross-peaks where one of the partners is difficult to assign univocally (overlapping of the signals), averaging on the sum of the distances involving the possible nuclei was chosen.

The topology file used here in X-PLOR 3.1 was derived from CHARMm's polysaccharideh.rtf in the Molecular Simulations package [27]. It was modified to incorporate specific atom types for cyclodextrin. Particularly, three different types of carbons were introduced for the glucose unit (first type: C-1, C-3, C-5, second type: C-2, C-4, third type: C-6) and three types of oxygen allowed the differentiation of O-4 from O-5 and from O-2, O-3, O-6. To maintain the $^4C_1$ conformation of the glucose units and the global torus-shape of the cavity during the high temperature dynamics, improper angles were added. The geometrical parameter values were derived from crystallographic data

[28]. The force constants were arbitrarily set to high values for the bond and for the improper terms in order to ensure covalent geometry.

(ii) Free molecular dynamics simulation in a box of water. In order to take into account the solvent explicitly, a molecular dynamic run was carried out in a waterbox TIP3P [29] of dimension $26.2 \times 26.2 \times 22.2$ Å^3 with periodic boundary conditions [30]. An integration time step of 1 fs in the Verlet algorithm was used to compute the trajectory. An invariant dielectric constant value equal to 1 was used during all the simulation. Distance cutoff on the non-bonded interactions was set to 11 Å. A 5–9 Å switching function was applied for the electrostatic and van der Waals interactions. The cyclodextrin was placed in a box of equilibrated water molecules, and all the solvent molecules overlapping the solute were removed. The number of remaining water molecules was 429, which was checked to correspond to the density of bulk water at 25°C. Heating to 300 K was realized in a first step of 3 ps. The temperature was equilibrated during 6 ps by rescaling the velocities. Then 257 ps of productive run was done. Monitoring the temperature and the total energy shows that the last 250 ps were stable. During the productive run, energies and coordinates were recorded every 1000 steps (1 ps). Update of the non-bond and image lists was realized every 5 steps (5 fs). All these molecular dynamics calculations in the waterbox were carried out with the CHARMm22 software [31]. The force field used is included in the Molecular Simulation package [27] and derives from the work of Brady [32].

(iii) Simulation of ring current effects. The chemical shift variations induced by the aromatic ring on the signals of the internal CD protons were estimated using a simplified law of the form $\Delta\delta_i = K(1-3\ \cos^2\theta_i)r_i^{-3}$ [33], where K is the dipolar strength depending on the aromatic ring, r_i the distance between the proton under consideration and the centre of the aromatic ring in Å and θ_i the angle between the vector r_i and the normal to the aromatic plane. As K is unknown, a constant scaling factor was applied in order to fit the experimental values.

3. Results and discussion

^{1}H NMR assignment strategy.—At first sight, the magnetic inequivalence between the protons of the different units indicates that the motion of the aromatic ring is restrained, at least according to the NMR time scale. This feature enables the complete assignment of all the proton signals which is necessary for a detailed analysis of the position of the aromatic ring and of the deformations of the cavity. Nevertheless, due to the extreme spectral overcrowding, it represents a very difficult task. Moreover the similarity in the chemical constitution of all the osidic units renders necessary a detailed stepwise analysis of the spin systems. Intra-unit assignment was performed using 2D semi-soft COSY and Relay [34,35]. In these experiments, H-1 protons were excited by a 270° gaussian pulse [36] covering selectively the anomeric region of the spectrum. Their magnetization was then transferred to H-2 in COSY, to H-2 and H-3 in the one-step Relay experiment, and so on. By this method we measured with a very high accuracy (estimated at 10^{-3} ppm) the chemical shift values for protons H-1, H-2, H-3, H-4 and H-5 of each unit (Table 1). Then a semi-soft TOCSY experiment was used to check

Table 1
Chemical shift values of H-1 to H-6′ for all CD units relative to TSP at 278 K. No stereospecific assignment has been done: H-6 and H-6′ are labelled according to their respective chemical shifts

Units	H-1	H-2	H-3	H-4	H-5	H-6	H-6′
I	5.009	3.583	3.768	3.296	3.355	3.226	3.571
II	5.143	3.717	4.083	3.666	3.980	4.051	4.163
III	5.049	3.650	3.834	3.581	3.688	3.917	3.982
IV	4.995	3.569	3.695	3.484	3.282	3.570	3.720
V	5.112	3.662	3.903	3.569	3.837	3.392	3.680
VI	5.092	3.717	4.174	3.675	3.949	4.000	4.070
VII	5.051	3.650	3.830	3.390	3.688	3.392	3.830

these assignments. At this point, the H-6 and H-6′ were not safety assigned due to non-efficient H-5 → H-6 or H-5 → H-6′ transfer in conjunction with the problems inherent to relaxation and spectral overcrowding. Their chemical shifts were determined using a triple-quantum filtered COSY experiment (with pulsed gradients for coherence selection [37]).

Sequencing of the cyclodextrin rim was achieved by a 2D semi-soft NOESY experiment exhibiting H-1–H-4 dipolar interactions between consecutive units. The following task consisted in determining the unit carrying the peptidic arm. ^{1}H NMR experiments in light water were performed in order to point out interactions through bonds between the amide proton and H-6, H-6′ of the substituted unit. A 1D semi-soft COSY with Jump–Return technique [20] used magnetization transfer between these resonances to provide their chemical shift values with a high accuracy (Fig. 1). The corresponding relay experiment intended to reveal the H-5 resonance was unsuccessful. To confirm these results, inverse heteronuclear ^{1}H–^{13}C correlation sequences were performed (not shown). HMQC [38] experiment followed by HMBC [39] confirmed the assignments made through ^{1}H homonuclear NMR. The former experiment shows connectivities between the carbons and protons directly bonded. The latter exhibits long-range interactions, mainly ^{2}J and ^{3}J. This allowed us to confirm the sequencing through H-1–C-4 interactions, and to find independently the substituted glucose unit after identification of the H-6–C-4 cross-peaks.

A first insight on the chemical shift values of the CD protons leads to some remarks. The high similarity for the proton chemical shifts between units II and VI, and between units III and VII suggests a preferential orientation of the aromatic ring according to the observed ring-current effects. However, no indication of a unique conformation can be assessed at this step, and an in-depth ^{1}H NMR analysis in addition with molecular modelling is necessary. Furthermore, according to the limited size of the cavity, it is necessary to distinguish accurately low differences in the distance constraints, especially between the CD and the tyrosine residue. These data are obtained by quantitative nOe evaluation.

Quantitative nOe evaluation.—The region of the NOESY contour maps concerning the interactions through space between the protons of the aromatic ring and some protons of the cavity is shown in Fig. 2. The build-up curve of the cross-relaxation

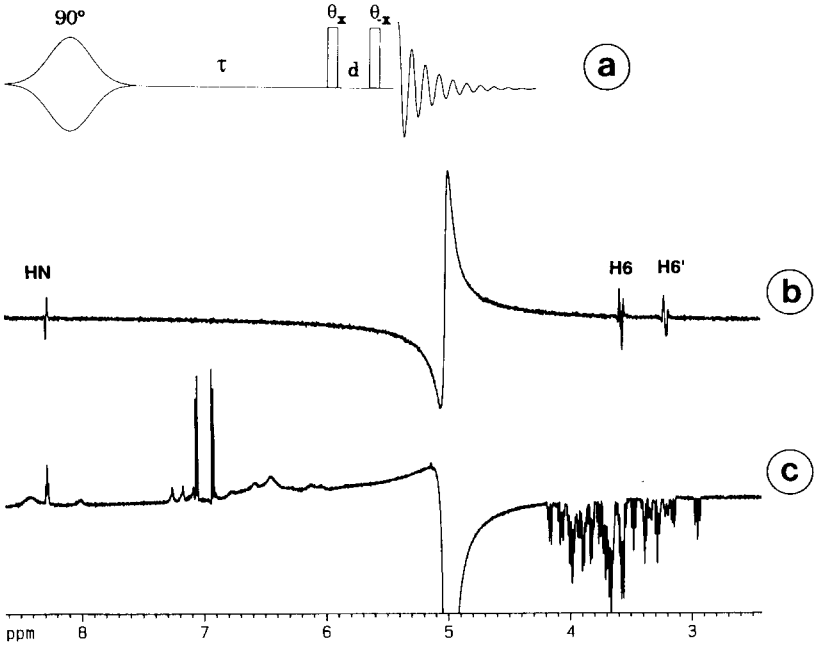


Fig. 1. Semi-soft 1D COSY with Jump–Return: (a) sequence, (b) sub-spectrum obtained after selective excitation of the amide proton. $\tau = 1/2J_{HN\text{-}H66'} = 30$ ms, $\theta = 45°$, $d = 160$ μs, (c) simple ^{1}H spectrum for comparison (Jump–Return technique).

between Hϵ of tyrosine and H-3 of unit VI is displayed as an example. The interactions between the amino acid residue and the cyclodextrin cavity are of major importance, as they represent the main NMR data for modelling the intra-molecular complex. For this reason, a "pseudo-3D" experiment was developed in ^{1}H NMR to check their existence. The sequence displayed in Fig. 3 consists in a 2D ROESY-TOCSY. One of the aromatic protons was selectively excited by a 270° gaussian pulse. Immediately after this selection, a low power spin-lock pulse was applied allowing dipolar interactions to occur in the rotating frame, i.e. magnetization to be transferred to protons H-3 or H-5 close in space. To ensure that no Hartmann–Hahn transfer between Hδ and Hϵ occurs during this step, the carrier frequency was shifted to the middle of the spectrum. The evolution delay followed, meaning that the diagonal was constituted by the signals of the H-3 and H-5 protons. The mixing period consisted in a short TOCSY composite spin-lock pulse (offset in the centre of the non-anomeric CD protons) designed to allow scalar magnetization transfer to the neighbouring spins, principally H-3 → H-4, H-2 and H-5 → H-4 (the transfer to H-6, H-6′ being inefficient). An important part of the contour plot obtained by this sequence after selective excitation of Hδ of tyrosine is displayed in Fig. 3. To illustrate the analysis of such a map, the diagonal peak corresponding to proton 3 of unit VI is circled and the row joining it to H-2 and H-4 is indicated by a dashed line. The presence of one or two cross-peaks at a given frequency in F1 is

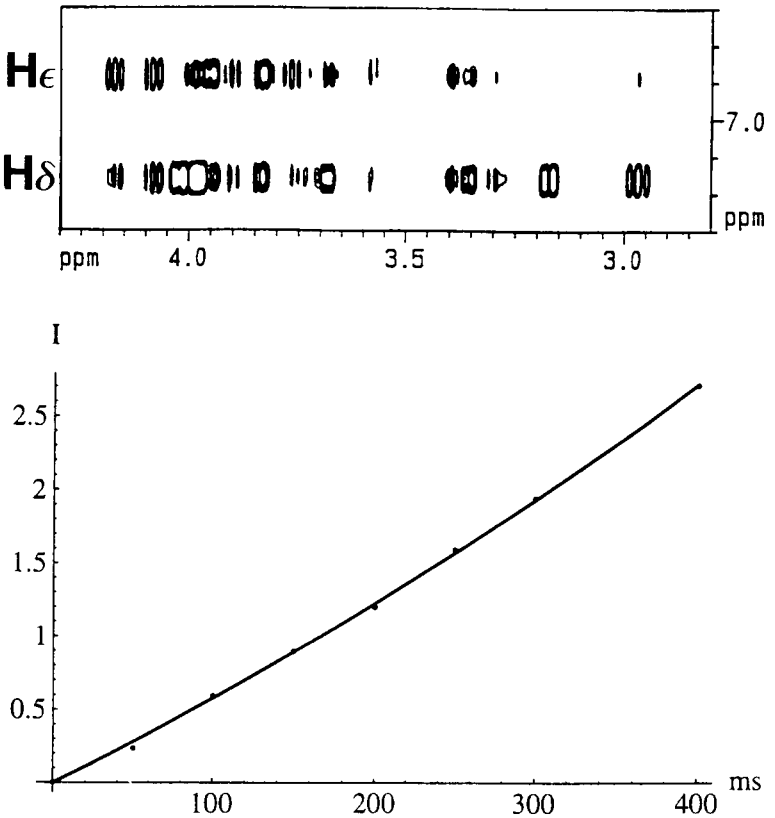


Fig. 2. Top: region of the NOESY contour plot corresponding to the interactions between aromatic protons and cyclodextrin; bottom: example of a build-up nOe curve: dipolar interaction $H\epsilon - {}^{VI}H3$.

generally adequate to identify with certainty the proton (on the diagonal) in dipolar interaction with $H\delta$.

Estimation of the local dynamics by NMR.—To explore local molecular dynamics, off-resonance steady-state magnetization measurements [22] were performed on the non-overlapped protons. This enabled the determination of the effective local correlation times, which can be used as an indication of the local dynamics. They are equal to the correlation times, if only dipolar interactions contribute to relaxation, and if in a local geometrical domain all the dipolar autocorrelation functions between proton pairs can be approximated by a single exponential with the same τ_c. No large discrepancy between the best-fit theoretical curves and experimental signal intensities and no geometrical discontinuity in the measures were observed for these protons. The effective local correlation time values could be classified into three groups: 0.55 ns for the aromatic protons, 0.60 ns for the anomeric protons and 0.85 ns for $H\beta$, $H\beta'$. The average value found is in agreement with the data in ^2H NMR [40]. For the aromatic and anomeric protons, the similar values of the effective local correlation times show that the

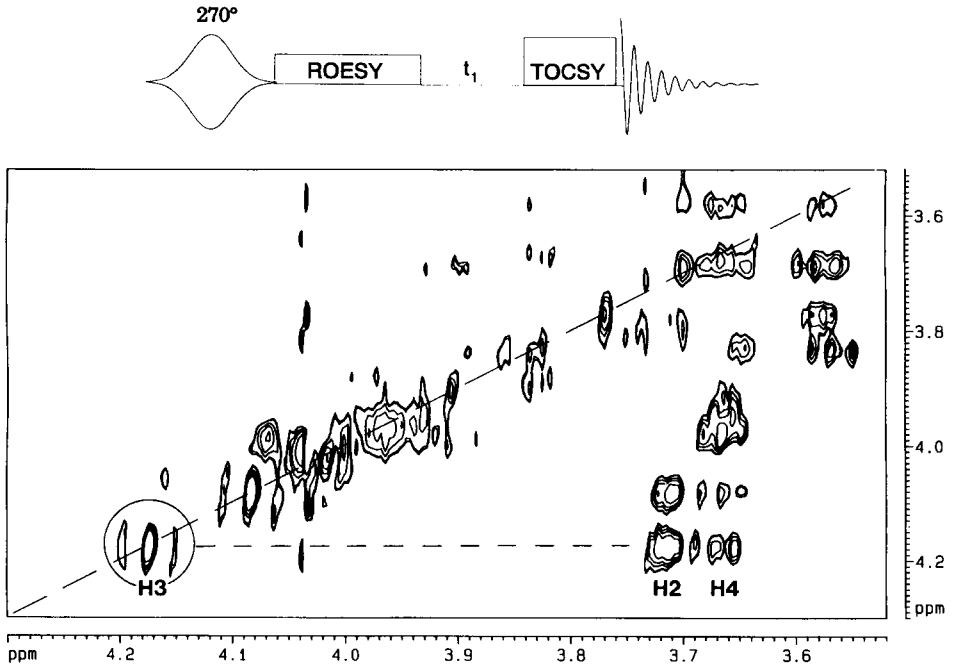


Fig. 3. Top: pulse sequence of the 2D ROESY-TOCSY experiment; bottom: partial contour plot obtained by this sequence after selective excitation of proton H δ.

interpretation of the cross-relaxation rates between aromatic protons and CD protons in terms of distances through a reference distance is a rather good approximation. However, the effective local correlation time values of the linker are larger than those of the CD and of the aromatic protons. It does not seem realistic for these protons to exhibit slower reorientational motion than the rest of the molecule. A more convenient explanation could be that relaxation mechanisms other than dipolar interactions are present, affecting the effective local correlation time. Fast chemical exchange—i.e. conformational exchange in the ms–μs time scale—could give a non-negligible contribution to proton relaxation. Therefore, interpretation of the cross-relaxation rates in terms of distances with the linker protons must be done carefully.

Three-dimensional structures.—Among the 50 structures generated by the SA procedure under nOe restraints (see Experimental section), 42 of them with no distance violation higher than 0.3 Å were kept. Most of the structures reveal only three distance violations, in the interval 0.2–0.3 Å. The covalent geometry was preserved: the average rms deviations from "ideal geometry" for bonds, angles and impropers were found to be weaker than 0.005 Å, 2° and 0.6° respectively.

Fig. 4 shows that the SA structures present little conformational heterogeneity. The position of the ring relative to the cyclodextrin units is the same in all the SA structures. For non-hydrogen atoms, the average rms deviation between the 42 structures and the

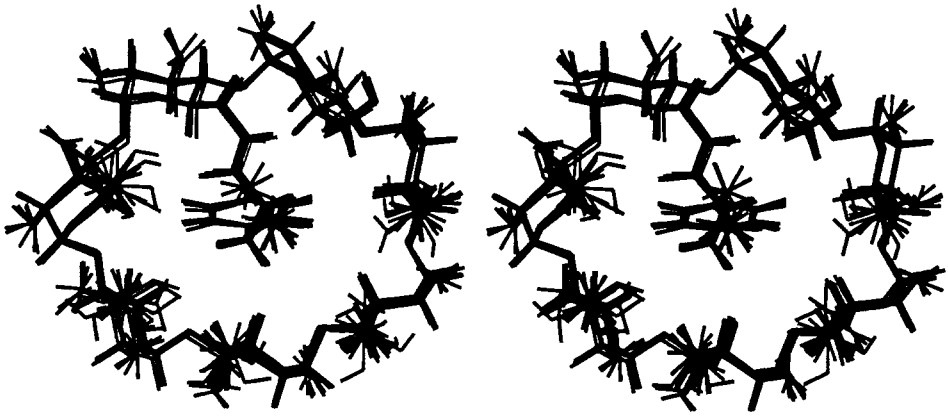


Fig. 4. Stereoviews from the primary side of the best-fit superposition of 10 SA structures of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose.

mean structure is equal to 0.3 Å. When removing O-6 of CD in the calculation, the average rms deviation decreases to 0.07 Å. It is distributed over the structure as follows: 0.06 Å for the cyclodextrin moiety, 0.07 Å for the aromatic ring, and 0.18 Å for the amino acyl link. This latter part of the structure appears less defined. However, it is not easy to conclude that it presents more fluctuations, as the lack of NMR data in this region due to experimental reasons could be the cause of such an apparent dispersion.

The first column of Table 2 presents the values of the different contributions in energy averaged over the SA structures, the second one the values for the average minimized SA structure. The low positive values of the van der Waals energy term indicate that the SA structures present no large steric effects between the included moiety and the CD rim. The values of the covalent energy terms reveal that the inclusion yields no large geometrical deformation of the CD.

In order to get information about the structure mobility, a molecular-dynamics simulation in a box of water was performed during 257 ps. One of the best SA structures was chosen as a starting structure. Over the last 250 ps, the temperature being stable ($T = 339\ \mathrm{K} \pm 10$), the average structure and the atomic fluctuations were calculated. Ten snapshots at 25-ps intervals are shown in Fig. 5. The mean structure was obtained by averaging the coordinates of 250 snapshots and applying 1000 steps of the adopted-basis Newton–Raphson minimization. The minimized mean MD structure was found very close to the average SA structure (Fig. 6): the rms deviation between them for the heavy atoms is 0.74 Å, while it decreases to 0.58 Å when the backbone atoms (heavy atoms without the O-6) are considered. Thus, during the MD dynamics simulation without any constraint, the structure does not deviate from that obtained by SA. In another MD simulation (not detailed here) started from a structure with the tyrosine ring arbitrarily rotated around Cγ–Cζ from its original position, the structure deviates and converges towards the SA structure. This ensures that the stability of the first MD trajectory was not the result of an incapacity to exit from a local minimum. In the simulation, the CD rim is not rigid enough to prevent any rotation of the tyrosine ring.

Table 2
Energy terms for the SA and MD structures

	SA averaged over the 42 structures (kcal/mol)	SA average minimized structure (kcal/mol)	MD averaged over the simulation (kcal/mol)	MD minimized mean structure (kcal/mol)
E total	61.7 ± 0.5	60.7	164.3 ± 13.1	−5.3
E bond	2.9 ± 0.1	2.7	70.0 ± 7.3	7.7
E angle	35.9 ± 0.9	34.3	105.6 ± 8.1	31.9
E dihedral	—	—	173.8 ± 2.9	167.3
E improper	5.9 ± 0.2	5.5	2.1 ± 1.0	0.2
E elec	—	—	−153.6 ± 8.3	−165.5
E vdw	16.2 ± 0.7	0.6	−33.7 ± 5.3	−47.1
E noe	35.9 ± 0.9	17.5	—	—

The last two columns of Table 2 present the values of the different contributions in energy averaged over the trajectory structures, and the corresponding values for the minimized mean MD structure. For the MD structures, the negative value of the van der Waals and electrostatic energy terms indicate that they contribute to stabilize the complex. The self-inclusion process involves an interaction energy between the two moieties of the complex which is equal to -48 ± 5 kcal/mol. This term includes the interaction between the aromatic ring and the CD (-13 ± 2 kcal/mol) and between the link and the CD (-35 ± 5 kcal/mol). These energies can be split into van der Waals and electrostatic terms. This decomposition shows that the CD-link interaction is mainly electrostatic (-29 ± 5 kcal/mol) with a small van der Waals component (-5 ± 1 kcal/mol). This electrostatic contribution is due to the NH_3^+ group of the amino acyl link. At the opposite end, the interaction energy between the aromatic ring and the CD is

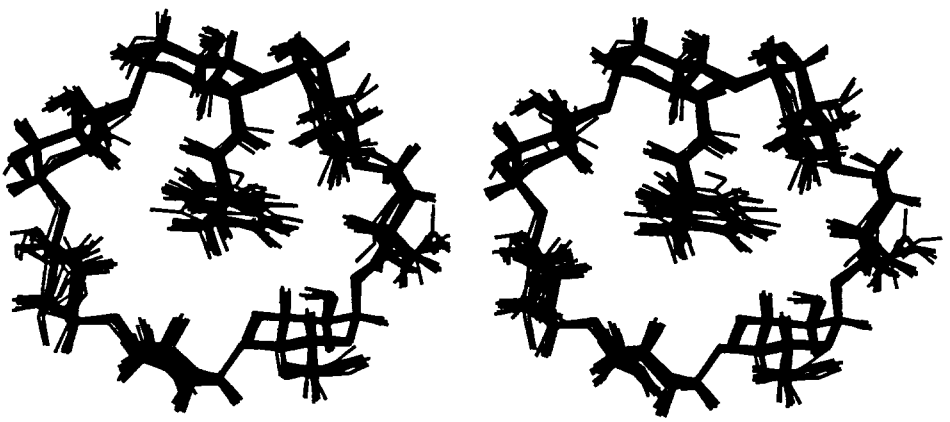


Fig. 5. Stereoviews from the primary side of the best-fit superposition of 10 MD structures of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose.

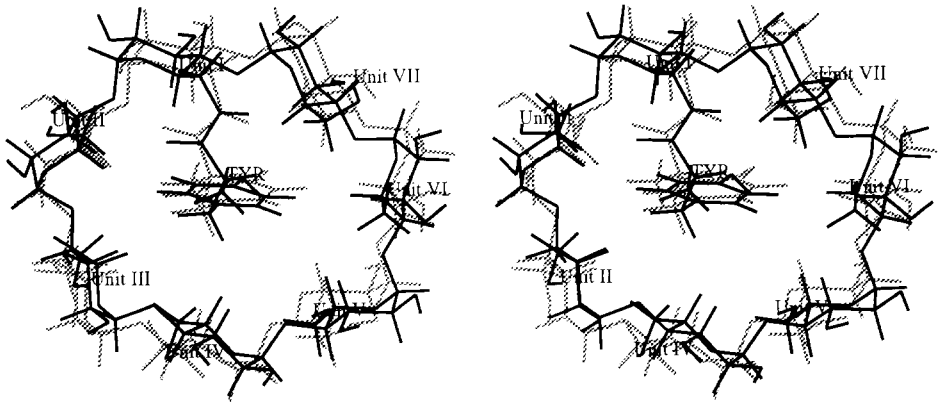


Fig. 6. Stereoviews of the best-fit superposition of the averaged SA (dark lines) and mean MD (grey lines) structures of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose.

mainly due to the van der Waals term (-11 ± 1 kcal/mol) with a small electrostatic component (-2 ± 2 kcal/mol). This van der Waals interaction energy is relatively strong, in the same order of magnitude as the corresponding value for a fully buried tyrosine ring in a protein [41].

The rms fluctuation averaged on the trajectory for all non-hydrogen atom is equal to 0.27 Å. For the atoms of the cyclodextrin part, the aromatic ring, and the amino acyl link the average fluctuations are equal to 0.25, 0.38, and 0.38 Å, respectively. These are larger than the corresponding values obtained for the SA structures. It shows that the flexibility of the structure is larger than indicated by a crude analysis of the rms deviation. The low values of the average fluctuations indicate that the self inclusion complex is rather rigid. All the parts of the structure present roughly the same rigidity except for the tyrosine part.

Back-calculation of experimental NMR data.—One of the main previous conclusions refers to the aromatic ring mobility which appears from the MD as being not larger than the mobility of the CD part. This is in agreement with the effective local correlation time values measured by NMR, as no significant difference in the dynamics of the aromatic ring and the cavity was pointed out.

In order to test the agreement of the molecular dynamics with the NMR data, the proton–proton distances associated to the nOe and the ring current effects were calculated. A set of 46 structures was obtained by selecting frames from the trajectory every 5 ps. The proton–proton distances were averaged in R^{-6} over the 46 structures and compared to the NMR distance constraints. Most of the differences between the experimental and simulation data are lower than the experimental error on the NMR distances which could be estimated to 20%. The only violation is equal to 0.3 Å. The ensemble of structures generated by MD is thus in agreement with these NMR data. From the 46 structures selected in the trajectory, all the proton–proton distances averaged in R^{-6} inferior to 3.5 Å were listed. Only one expected cross peak is missing in the NOESY map. This concerns a spatial proximity between HN of the peptidic arm

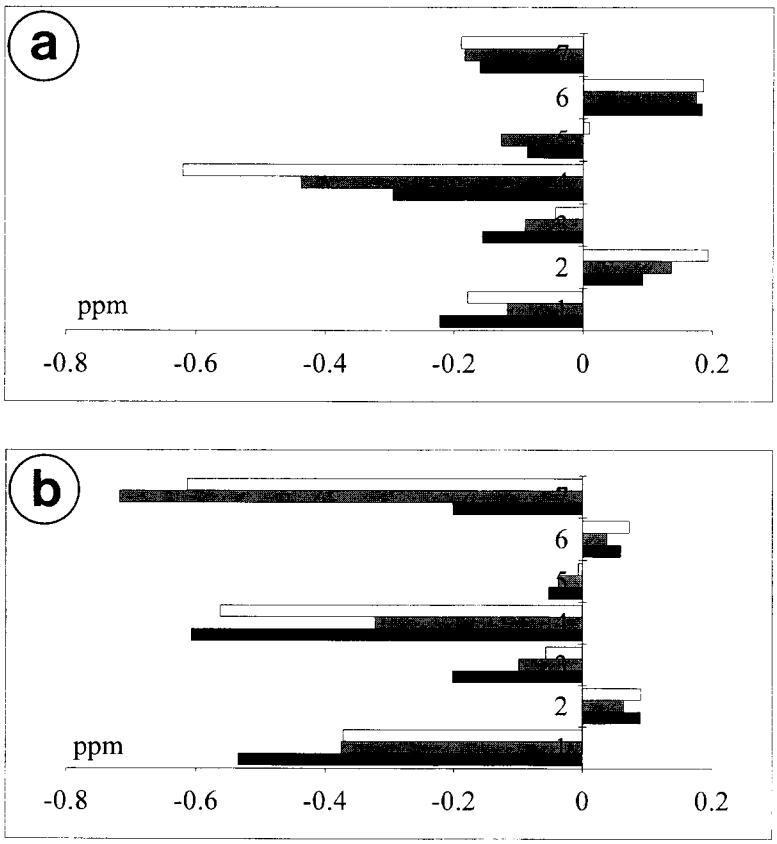


Fig. 7. Chemical shift variations induced by ring current effects of the aromatic ring on (a) protons H-3 and (b) protons H-5. Black rectangles: experimental data; white rectangles: values predicted from MD structures; grey rectangles: values predicted from SA structures.

and H-5 of unit VII. The absence of this cross peak could be due to a conformational exchange—previously mentioned—involving the peptidic arm in the ms–μs time scale. Neither the analysis of the SA structures, nor that of the MD structure gives evidence of such a movement. In SA calculation, the derived proton–proton distances are not sufficiently precise and their interpretation does not take into account the movement. For this reason, the ensemble of SA structures is not properly related to all accessible conformations in solution. In the MD simulation, the trajectory is not long enough to include such time scale motions.

A more sensitive method to evaluate the agreement with the NMR data is to calculate the ring current effects induced by the aromatic ring on the H-3 and H-5 protons located inside the cavity. The resulting chemical-shift variations have been estimated using the simplified law given in the Experimental section. A comparison between the calculated chemical-shift variations and the corresponding experimental data is given in Fig. 7. The

former were averaged over all the SA structures in a first approach, and over the 46 structures obtained from the trajectory in a second approach. The rms deviations between calculated and experimental values are, considering H-3, 0.15 and 0.08 ppm for the SA and MD structures, respectively, whereas for H-5 they reach 0.19 and 0.25 ppm. However, this larger deviation can arise from the anisotropy of the carbonyl which could interfere on the observed effect. Indeed, when unit VII is not taken into account, the rms

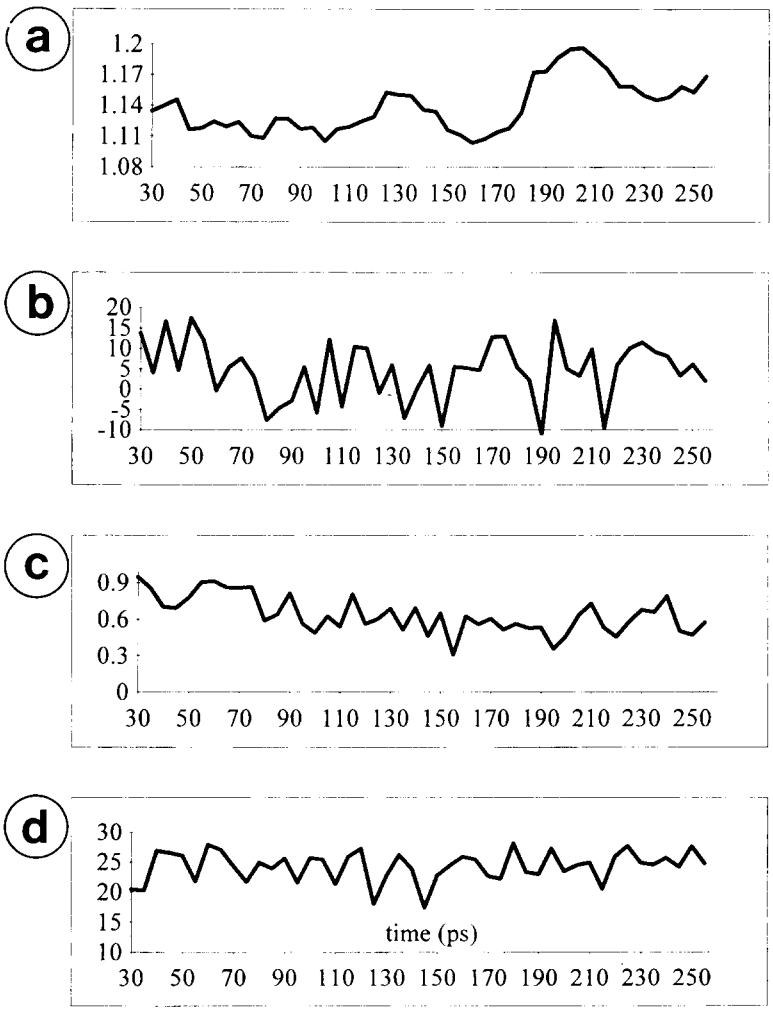


Fig. 8. Characteristic time series for the self-inclusion complex: (a) ratio of the large diameter over the small one for the ellipse formed by the O-4 of CD; (b) angle between the aromatic plane and the vector ^{III}O-4–^{VI}O-4; (c) distance between the aromatic centre and the centre of the O-4; (d) angle between the aromatic plane and the normal of the least-square plane of the O-4.

deviations fall to 0.07 and 0.12 ppm. It indicates an almost perfect matching of the calculated values with the experimental ones. This leads to the conclusion that both the orientation of the Tyr ring and the distances between the aromatic ring and the internal protons of the CD are well represented in these two approaches. Furthermore this good result also shows that even an oversimplified dipolar law can accurately represent the experimental observations when the interactions occur on similar units in a restrained space.

Structural and dynamical analysis of the molecular dynamics trajectory.—(i) Structure and dynamics of the self-included complex. Considering the previous results, the MD structures reproduce the NMR data with a good accuracy. It is thus possible to derive the structural and dynamic properties of the compound by analysis of the trajectory. In order to characterize the structure of the self-inclusion complex, we have calculated the four following time series which are shown in Fig. 8:

1. The ratio of large axis over small axis length for the ellipse formed by the interglycosidic O-4 atoms of CD (equal to 1.14 ± 0.06),
2. The angle between the aromatic plane and the vector ^{III}O-4–^{VI}O-4 (equal to $4.6° \pm 7.2$),
3. The distance between the centre of the aromatic ring and the centre of the O-4 (equal to 0.64 Å $\pm$ 0.15),
4. The tilt angle between the aromatic plane and the normal of the least-square plane of the O-4 (equal to $24.27° \pm 2.5$).

These time series present relatively small fluctuations and do not seem correlated. Their average values are discussed in the following parts. The CD rim does not present any significant twist, all O-4 nearly defining a plane. These atoms form an ellipse for which the ratio big axis length over small is 1.14. This value is quite large compared to the deformation of the cavity for non-covalent α-CD complexes with several guests [42]. Note, however, that Harata et al. have found the same value as us in an X-ray structure of the non-covalent β-cyclodextrin–benzyl alcohol complex [43].

The two distances ^{III}O-4–^{VI}O-4 and ^{V}O-4–^{I}O-4 are respectively the longest (10.4 Å) and the shortest (9.2 Å) O-4–O-4 distances between units i and $i+3$. The value of the angle between the aromatic plane and the vector ^{III}O-4–^{VI}O-4 is close to zero, meaning that the former is nearly parallel to the big axis and roughly perpendicular to the link.

The presence at the extremity of the big axis of the ellipse of numerous van der Waals contacts between tyrosine and the CD units suggests that the CD deformation is induced by the aromatic ring. All 13 van der Waals contacts existing between Tyr protons and H-3, H-5 of CD concern units II, III, V and VI. Most of them involve H-5 (eight van der Waals contacts) due to the narrower hole of the primary side. The shortest contact at one side of the aromatic plane is Hδ–^{II}H-5 (2.21 ± 0.15 Å), and at the other side Hϵ'–H-3 and Hϵ'–^{VI}H-3 (2.51 ± 0.24, 2.55 ± 0.35 Å, respectively). The low level of the fluctuations of these distances indicates that the van der Waals contacts are quasi permanent, requiring that the CD adopts an elliptical shape.

The third parameter value indicates that the centre of mass of the tyrosine ring is close to the centre of the CD moiety defined from the O-4 atoms (distance equal to 0.64 Å $\pm$ 0.15). The aromatic ring is slightly shifted above the O-4 plane towards the narrow side of the cavity (0.35 Å), and translated closer to the unit VII (0.5 Å). In this position,

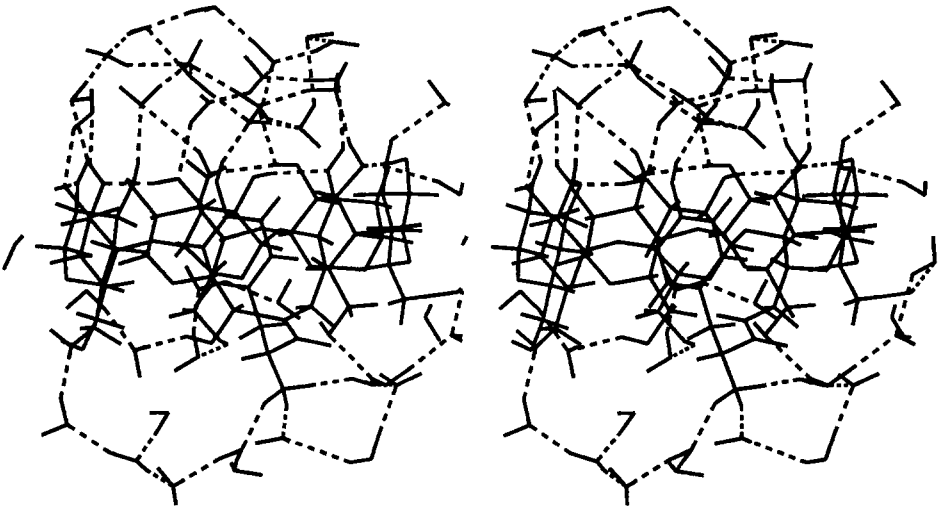


Fig. 9. Stereoview of the mean structure of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose with water molecules of the first shell. A display cutoff was applied beyond 10 Å from the centre of the solute. Hydrogen bonds are represented by dashed lines.

the hydroxyl group of Tyr is at the same level as OH-2 and OH-3 of CD. This location of the guest in the cavity, which is in agreement with previous results obtained by Inoue and co-workers [44,45], enables the host and guest hydroxyls to form all together an hydration shell which may favour complex solvation (see Fig. 9). The aromatic ring is not perpendicular to the O-4 plane, but forms an angle equal to $24.27° \pm 2.5$, in such a way that whereas Cβ remains centred, the hydroxyl group of Tyr is directed towards unit VII.

The minimized mean MD structure is very close to the X-ray structure of the non-covalent β-cyclodextrin–benzyl alcohol complex yet mentioned in ref. [43], as shown in Fig. 10: (i) the rms deviation between the CD backbone atoms is equal to 0.43 Å, (ii) the aromatic rings have roughly the same orientation (the difference between their tilt angles is 4°), (iii) the hydroxyls of the guests have the same location relative to the OH-2 and OH-3 of the CD. In our opinion, the slight difference in the tilt angle should allow the system to align in both cases the guest hydroxyl with the secondary hydroxyls of CD.

An analysis of the fluctuations on the atomic coordinates during the MD simulation leads to three separate groups of units according to their relative mobilities. Unit III presents the most important fluctuations (0.3 Å), units II, IV, V are slightly more constrained (fluctuations around 0.25 Å). Units I, VI, VII show the lower rms fluctuation (around 0.2 Å). All these results can be explained by the offset of the aromatic ring from the centre of the cavity towards unit VII, and by its tilt, which brings the Tyr ring nearer to units I, VI, VII than to II, III, IV, V. This is specially true when comparing for example the distances Cζ–H-5, which can be classified in increasing order as follows:

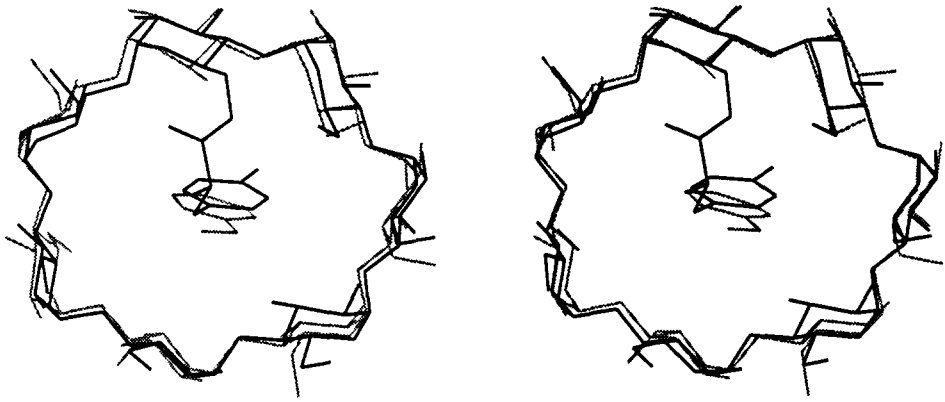


Fig. 10. Stereoview of the best-fit superposition of the minimized mean MD structure of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose (dark lines) and X-ray structure of the non-covalent cyclomaltoheptaose–benzyl alcohol complex (grey lines). The latter is available from the Cambridge Structural Database (REFCODE: DEBGOG).

unit VII < unit I < unit VI < unit V < unit IV < unit II < unit III, from $d = 3.3$ Å for unit VII to $d = 5.6$ Å for unit III.

One of the main questions in the study of host–guest complexes relates to the respective motions of the host and the guest. The NMR determination of the effective local correlation times suggests that the dynamics of the aromatic ring is restrained and that there is no significant difference in the dynamics of the "host" and the "guest". Analysis of the molecular dynamics simulation confirms this result and provides two explanations about this behaviour. Firstly, their respective size reveals that the latter tightly fits the former. As previously mentioned, the van der Waals contacts between the moieties are strong, making the guest not able to move independently of the host. Secondly, the peptidic link may also have a double role. Its length and intrinsic flexibility may place the guest in such a position in the CD rim that no motion is possible. The link may also present steric effects which prevent the motion of the aromatic ring relative to the CD.

(ii) Analysis of the hydrogen bonds. The self included complex exhibits intramolecular hydrogen bonds between oxygens of secondary hydroxyls O-2 and O-3, which can act simultaneously as hydrogen bonding donors and acceptors. We used as a criterion for hydrogen bonding that the hydrogen–acceptor atom distance has to be less than 2.4 Å and the angle donor–H–acceptor greater than 135°. Table 3 shows the percentage of occurrence of these intramolecular H-bonds. H-Bonds between these atoms and solvent molecules are also indicated, as it shows that hydroxyls not involved in such intramolecular bonds participate in a hydrogen bonding network with the water.

Two flip-flop phenomena of the type ${}^{i}\text{O-2–H} \cdots {}^{i+1}\text{O-3} \leftrightarrow {}^{i}\text{O-2} \cdots {}^{i+1}\text{HO-3}$ between units III–IV and VI–VII were observed during the simulation. Between glucose units I–II the hydrogen bond ${}^{\text{I}}\text{HO-2} \cdots {}^{\text{II}}\text{O-3}$ is present for 80%, and the reverse hydrogen bond ${}^{\text{I}}\text{O-2} \cdots {}^{\text{II}}\text{HO-3}$ is present for less than 5% (at the beginning of the

Table 3
Percentages of intramolecular hydrogen bonds and with the solvent involving the secondary hydroxyls. Only the percentages of occurrence superior to 5% are given. In parenthesis, the name of the atom involved in a H-bond with a water molecule is indicated. [*]: two H_2O molecules

% of H-Bonds between units:	Type $^{i}HO\text{-}2\cdots{}^{i+1}O\text{-}3$	Type $^{i}O\text{-}2\cdots{}^{i+1}HO\text{-}3$	% of H-Bonds with the solvent	
I–II	80.0			14.5 (^{I}O-2)
				36.7 (^{II}HO-3)
II–III	83.5			77.0 (^{III}HO-3)
III–IV	25.8	8.9	16.1 (^{IV}O-3)	48.8 (^{IV}HO-3)
IV–V		95.5	98.4 (^{IV}HO-2)	
			16.9 (^{V}O-3)	
V–VI		90.7	60.5 (^{V}HO-2)	
			58.5 (^{VI}O-3)	
VI–VII	43.1	51.2	50.4 (^{VI}HO-2)	21.4 (^{VI}O-2)
			40.7 + 18.5 [*] (^{VII}O-3)	15.3 (^{VII}HO-3)
VII–I	97.6			79.0 (^{VII}O-2)
				99.6 (^{I}HO-3)

simulation). The four other intramolecular interglucose hydrogen bonds are unidirectional. For units IV–V and V–VI the hydrogen bond is in one sense, for VII–I and I–II the reverse.

The flip-flop hydrogen-bonding phenomenon was extensively studied by MD simulation [46–48]. Many complex patterns of the hydrogen bonding network which are found here correspond to previous descriptions. This contributes further to validate the present simulation.

Analysis of the first-shell water molecules shows that the hydroxyl rim and the NH_3^+ present hydrogen bonds to water molecules (Fig. 9). These bonded water molecules present a lower fluctuation than the bulk water. The tyrosine ring fills completely the CD so that no water molecule is present inside the CD cavity. The NH_3^+ ion is strongly bonded to three water molecules which form an hydration shell stabilizing this ion over the narrow side of the cavity. At the other side of the cavity, the hydroxyl of tyrosine is hydrated by one water molecule, but this interaction occurred only during half the time of the simulation. As indicated previously, this hydroxyl and the secondary hydroxyls of the CD are bonded on average to 10 water molecules which form an hydration shell at the large side of the cavity.

4. Conclusion

We have shown that even in the case of an intense spectral overcrowding (43 proton signals are present in a 0.8-ppm range) it is possible to obtain quantitative NMR data which can be safely used in molecular modelling to determine the 3D structure. Through dedicated soft experiments like the pseudo-3D ROESY-TOCSY sequence, a complete and univocal assignment has been performed. Then NOESY experiments provided H–H

distance constraints as accurately as possible. [1] Although a relatively low number of "host–guest" constraints has been obtained, their use in the simulated annealing procedure affords a unique and reliable family of structures. This is certainly due to the reduced space of the interior of the cavity. The chemical-shift variations induced by an aromatic guest on the internal protons of the host allow precise checking of both the position of the former and the deformation of the cavity. It may be useful in the future to use them directly as constraints in the simulation procedure. Both the SA and MD structures are in good agreement with the experimental NMR data. [2] Especially, the ring current effects are well reproduced.

The results show that the tyrosine ring is totally blocked inside the cavity and does not experience any independent motion relative to CD. The aromatic part of the peptidic side chain fully fills the cavity and induces a deformation of the CD rim. It takes an elliptical shape with its large diameter quasi-parallel to the aromatic plane in order to receive it. The compound is rigid with no part really flexible. The rigidity is attributable not only to the respective sizes of the host and the guest, but mainly to the presence of the arm which links the guest to the host. Steric effects involving the arm may influence their respective orientation and their relative motion. The ^{1}H NMR spectrum of the non-covalent complex β-cyclodextrin–L-tyrosine has been performed in the same conditions as in the present study. It exhibits for the host and guest signals a non-perturbed symmetry. This preliminary result indicates that the tyrosine rotates isotropically in the cavity.

The global rigidity of the self-inclusion complex is in contrast with the results found for pure non-covalent host–guest complexes, which are very flexible and present a reciprocal conformational adaptation of the host to the guest [50]. However, a simple comparison is not so easy as even a simple chemical modification of the CD can affect the dynamics of the guest by a molecular-recognition process [7]. The presence of a substituent grafted on the CD restricts the relative host–guest motion as it bounds the available conformational space of the guest. To allow a better understanding of this effect, we are studying the inclusion phenomenon in three systems. The first one consists of βCD substituted by L-tyrosine at C-3, which enables the exploration of the influence of the size on the inclusion side. The second one is the non-covalent complex constituted by 6′-amino 6′-deoxycyclomaltoheptaose in equimolar solution with L-tyrosine. Finally, in order to estimate the influence of the amide bond on the geometrical and dynamical properties of the complex, the study of an analog with a different host–guest link is under progress.

[1] Since we performed the NMR experiments described in this paper, we have developed a new pulse sequence called off-resonance ROESY [49]. It overcomes the problem of longitudinal cross-relaxation vanishing, common in the study of medium-sized molecules at ambient temperature, which obliged us to work at low temperature.

[2] The atomic coordinates of the average SA and MD structures, the list of the 21 host–guest nOe distances constraints inferior to 4 Å compared to the values found in SA and MD approaches, the interglycosidic angles, and the dihedral angles defining the aminoacyl linkage are available upon request from the authors.

Acknowledgements

The authors thank Drs B. Perly and F. Djedaïni-Pilard for providing the sample of 6-deoxy-6-L-tyrosinylamidocyclomaltoheptaose, and Dr J. Smith for helpful discussion about this manuscript.

References

[1] M.L. Bender and M. Komiyama, *Cyclodextrin Chemistry*, Springer, Berlin, 1978.
[2] D. Duchêne (Ed.), *New Trends in Cyclodextrins and Derivatives*, Editions de Santé, Paris, 1989.
[3] J. Szejtli, *Cyclodextrin Technology*, Kluwer Academic Publishers, Dordrecht, 1988.
[4] I. Tabushi, Y. Kuroda, and T. Mizutani, *J. Am. Chem. Soc.*, 108 (1986) 4514–4518.
[5] A.P. Croft and R.A. Bartsch, *Tetrahedron*, 39 (1983) 1417–1474.
[6] A. Ueno, Y. Tomita, and T. Osa, *J. Chem. Soc., Chem. Commun.*, (1983) 976–977.
[7] I. Tabushi, Y. Kuroda, and M. Yamada, *Tetrahedron Lett.*, 29 (1988) 1413–1416.
[8] R. Breslow and A.W. Czarnik, *J. Am. Chem. Soc.*, 105 (1983) 1390–1391.
[9] K.R. Rao, T.N. Srinivasan, N. Bhanumathi, and P.B. Sattur, *J. Chem. Soc., Chem. Commun.*, (1990) 10–11.
[10] K. Takahashi, Y. Ohtsuka, and K. Hattori, *Chem. Lett.*, (1990) 2227–2230.
[11] F. Djedaïni-Pilard, J. Desalos, and B. Perly, *Tetrahedron Lett.*, 34 (1993) 2457–2460.
[12] W. Saka, Y. Yamamoto, Y. Inoue, R. Chûjô, K. Takahashi, and K. Hattori, *Bull. Chem. Soc. Jpn.*, 63 (1990) 3175–3182.
[13] H. Parrot-Lopez, H. Galons, A.W. Coleman, F. Djedaïni, N. Keller, and B. Perly, *Tetrahedron: Asymmetry*, 1 (1990) 367–370.
[14] K. Takahashi, Y. Ohtsuka, S. Nakada, and K. Hattori, *J. Inclusion Phenom. Mol. Recogn. Chem.*, 10 (1991) 63–68.
[15] K. Takahashi, *Bull. Chem. Soc. Jpn.*, 66 (1993) 550–554.
[16] K. Takahashi and K. Hattori, *Supramol. Chem.*, 2 (1993) 305–308.
[17] F. Djedaïni-Pilard, N. Azaroual-Bellanger, M. Gosnat, D. Vernet, and B. Perly, *J. Chem. Soc., Perkin Trans. 2*, (1995) 723–730.
[18] D. Marion and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, 113 (1983) 967–974.
[19] D. Neuhaus and M. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, New York, 1989.
[20] P. Plateau and M. Guéron, *J. Am. Chem. Soc.*, 104 (1982) 7310–7311.
[21] H. Desvaux, PhD Thesis, Université de Paris VI, 1993.
[22] H. Desvaux and M. Goldman, *Mol. Phys.*, 81 (1994) 955–974.
[23] A. Bax and R. Freeman, *J. Am. Chem. Soc.*, 104 (1982) 1099–1100.
[24] I. Tvaroska, M. Hricovini, and E. Petrakova, *Carbohydr. Res.*, 189 (1989) 359–362.
[25] A.T. Brünger, *X-PLOR Version 3.1, a System for X-ray Crystallography and NMR*, 1992.
[26] M. Nilges, G.M. Clore, and A.M. Gronenborn, *FEBS Lett.*, 239 (1988) 129–136.
[27] Molecular Simulations Inc., 16 New England Executive Park, Burlington, MA, USA.
[28] K. Lindner and W. Saenger, *Carbohydr. Res.*, 99 (1982) 103–115.
[29] W.L. Jorgensen, *J. Am. Chem. Soc.*, 103 (1981) 335–340.
[30] C.L. Brooks III, M. Karplus, and B.M. Pettit, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics. Advances in Chemical Physics*, Vol. LXXI, Wiley, New York, 1988.
[31] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, 4 (1983) 187–217.
[32] S.N. Ha, A. Giammona, M. Field, and J.W. Brady, *Carbohydr. Res.*, 180 (1988) 207–221.
[33] C.E. Johnson Jr. and F.A. Bovey, *J. Chem. Phys.*, 29 (1958) 1012–1014.
[34] P. Berthault, F. Djedaïni, and B. Perly, *J. Magn. Reson.*, 91 (1991) 102–112.

[35] P. Berthault, H. Desvaux, and B. Perly, *Magn. Reson. Chem.*, 31 (1993) 259–265.
[36] L. Emsley and G. Bodenhausen, *J. Magn. Reson.*, 82 (1989) 211–221.
[37] A.L. Davis, E.D. Laue, J. Keeler, D. Moskau, and J. Lohman, *J. Magn. Reson.*, 94 (1991) 637–644.
[38] L. Müller, *J. Am. Chem. Soc.*, 101 (1979) 4481–4484.
[39] A. Bax and M.F. Summers, *J. Am. Chem. Soc.*, 108 (1986) 2093–2094.
[40] N. Bellanger, PhD Thesis, Université de Paris VI, 1993.
[41] G.I. Makhatadze and P.L. Privalov, *J. Mol. Biol.*, 232 (1993) 639–659.
[42] S.P. van Helden, M.J. van Drooge, A.J. Claessens, A.C.A. Jansen, and L.H.M. Janssen, *Carbohydr. Res.*, 215 (1991) 251–260.
[43] K. Harata, K. Uekama, M. Otagiri, F. Hirayama, and Y. Ohtani, *Bull. Chem. Soc. Jpn.*, 58 (1985) 1234–1238.
[44] Y. Inoue, *Ann. Rep. NMR Spectroscopy*, 27 (1993) 59–99.
[45] Y. Inoue, T. Okuda, Y. Miyata, and R. Chûjô, *Carbohydr. Res.*, 125 (1984) 65–76.
[46] C. Betzel, W. Saenger, B.E. Hingerty, and G.M. Brown, *J. Am. Chem. Soc.*, 106 (1984) 7545–7557.
[47] J.E.H. Koehler, W. Saenger, and W.F. van Gunsteren, *J. Biomol. Struct. Dyn.*, 6 (1988) 181–198.
[48] J.E.H. Koehler, W. Saenger, and W.F. van Gunsteren, *Eur. Biophys. J.*, 16 (1988) 153–168.
[49] H. Desvaux, P. Berthault, N. Birlirakis, and M. Goldman, *J. Magn. Reson., Ser. A*, 108 (1994) 219–229.
[50] S.P. van Helden, B.P. van Eijck, and L.H.M. Janssen, *J. Biomol. Struct. Dyn.*, 9 (1992) 1269–1283.