Molecular recognition of *N*-protected dipeptides by pseudopeptidic macrocycles: a comparative study of the supramolecular complexes by ESI-MS and NMR†

Ignacio Alfonso,*[a] Michael Bolte,[b] Miriam Bru,[c] M. Isabel Burguete,[c] Santiago V. Luis*[c] and Cristian Vicent[d]

Received 27th November 2009, Accepted 21st December 2009
First published as an Advance Article on the web 25th January 2010
DOI: 10.1039/b924981h

The molecular recognition properties of pseudopeptidic macrocycles have been studied by ESI-MS and NMR spectroscopy, as highly complementary experimental techniques in solution and in the gas phase. We used ESI-MS competition experiments for the high throughput screening of the supramolecular interaction between four macrocyclic receptors and different peptide-like substrates in solution, rendering the best-fitted host–guest pairs. Further insights on the non-covalent recognition process in the gas-phase were obtained through collision induced dissociation (CID) experiments. Solution studies using NMR spectroscopy (^{1}H NMR titrations, NOESY and DOSY) were carried out to prove the validity of ESI-MS as a high-throughput screening method for studying the molecular recognition of the investigated pseudopeptidic macrocycles. A clear selectivity for *N*-protected dipeptides over *N*-protected amino acids, and a slight preference for dipeptides bearing aromatic side chains were observed. On the basis of the results obtained from this approach, a mode of binding has been proposed.

Introduction

Understanding and controlling peptide–peptide non-covalent interactions is a major challenge in supramolecular[1] and biological chemistry.[2] These interactions are ultimately responsible for the protein–protein contacts,[3] which are key physico-chemical processes for the regulation of fundamental biological functions.[4] The modulation of these interactions is biologically very important, with promising therapeutic applications, through either promotion or inhibition of the processes.[5] Thus, some illnesses have their fundamental origin in the self-aggregation of peptide oligomers, leading to fibrillar precipitates which finally produce the pathogenic structures.[6] This is the case of some neurological fatal diseases such as Parkinson's disease,[7] Huntington's disease[8] and Alzheimer's disease;[9] or metabolic chronic disorders like some types of diabetes.[10] On the other hand, the mechanism of action of some antibiotics, like vancomycin, relies on the specific molecular recognition of the D-Ala-D-Ala carboxylic terminus of the bacterial cell wall peptidoglycan.[11] Another important example is the binding of Arg-Gly-Asp sequence to promote cell adhesion and growing.[12] In the recent literature, there are several examples of synthetic receptors for short peptides,[13] able to selectively bind given sequences (such as D-Ala-D-Ala,[14] Phe-Phe[15] or Arg-Gly-Asp[16]) in different environments. Usually, this selectivity is obtained by a synergic action of non-covalent contacts, such as coordinative, electrostatic, H-bonding, π–π, hydrophobic or steric interactions.[17] Therefore, the detailed structural information about the specific host–guest species is highly valuable for optimization of the non-covalent interactions. Within this context, several experimental techniques have been used for the study and characterization of the supramolecular complexes formed upon the binding phenomena, although there is no general approach in this regard.[18] The selection of the most suitable technique depends on the receptor structure and on the strength of the complexes formed. Regarding that, the combined use of several techniques is advisable in order to compare results from different sources and, if possible, to extract conclusions in different environments and experimental conditions. Among others, Nuclear Magnetic Resonance is, by far, the most powerful technique for obtaining deep structural information.[19] Very important features like conformation, binding epitopes, internuclear distances, or size and shape of the complexes formed in solution can be determined with specific NMR-based experiments. On the contrary, NMR has relatively low sensitivity, being sample and time consuming.

Another interesting technique, less often used in the field of supramolecular chemistry, is mass spectrometry.[20] Soft ionization mass spectrometric techniques have appeared in the last decade as a major breakthrough for the study of host–guest complexes. In particular, electrospray ionization mass spectrometry

[a]*Departamento de Química Biológica y Modelización Molecular, Instituto de Química Avanzada de Cataluña, Consejo Superior de Investigaciones Científicas, IQAC/CSIC, C/Jordi Girona18-26, E-08034, Barcelona, Spain. E-mail: ignacio.alfonso@iqac.csic.es; Fax: +34 932045904; Tel: +34 934006100 ext: 1381*
[b]*Institut für Anorganische Chemie, J. W. Goethe-Universität, Frankfurt, Max-von-Laue-Str. 7, 60438, Frankfurt/Main, Germany*
[c]*Departamento de Química Inorgánica y Orgánica, Universidad Jaume I, Avenida Sos Baynat s/n, E-12071, Castellón, Spain. E-mail: luiss@qio.uji.es; Fax: +34 964728214; Tel: +34 964728239*
[d]*Serveis Centrals d'Instrumentació Científica, Universitat Jaume I, Castellón, Spain*

† Electronic supplementary information (ESI) available: Detailed experimental procedures, additional ESI-TOF and CID mass spectra, full 1D and 2D (gCOSY, TOCSY, NOESY and ROESY) NMR spectra, NMR titration data and procedures, X-ray crystallographic table and CIF file for *Z*-AF-OH, Cartesian coordinates (*XYZ*) for the minimized structure and HYDRONMR output file. CCDC reference number 739735. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/b924981h

(ESI-MS) has proved to be a very powerful tool to characterize non-covalent complexes as illustrated for several hosts such as crown-ethers,[21] calixarenes,[22] cyclodextrins,[23] cucurbit[*n*]urils[24] or resorcinarenes.[25] The main advantages of MS-based techniques are fast performance, low sample consumption and the possibility to use relatively complicated mixtures of compounds, which make mass spectrometry very interesting for high throughput screening.[26] Besides, through ESI-MS, it is possible to define the stoichiometry and the topology of the interacting species as well as to correlate peak intensities directly from ESI mass spectra to quantitative solution-phase data (relative or absolute binding affinities of the non-covalent complexes) according to two experimental approaches, the so-called titration and competition experiments.[26] In this latter issue, limitations based on such correlation are usually system dependent, and care should be taken when trying to quantitatively correlate observed gas phase ion abundances with solution phase concentrations.[27,28] For this reason, control experiments comparing experimental data from ESI-MS and other solution-phase techniques (typically NMR) are mandatory to validate the ESI-MS approach as a true diagnostic of the solution speciation which is the primary focus of the present work.

Therefore, we believe that both techniques (NMR and ESI-MS) are complementary in essence, since the advantages of each of them would solve the disadvantages of the other. We have used this rationale for studying the abilities of synthetic pseudopeptidic macrocycles[29] (see Chart 1) as molecular receptors for peptide-like substrates either in solution or in the gas-phase.[30] The chemical structures of the macrocycles were selected in order to map two variables: the nature of the side chain (aromatic/aliphatic) and the spacer between the amino amide moieties (ethylene/cyclohexane-1,2-diyl).

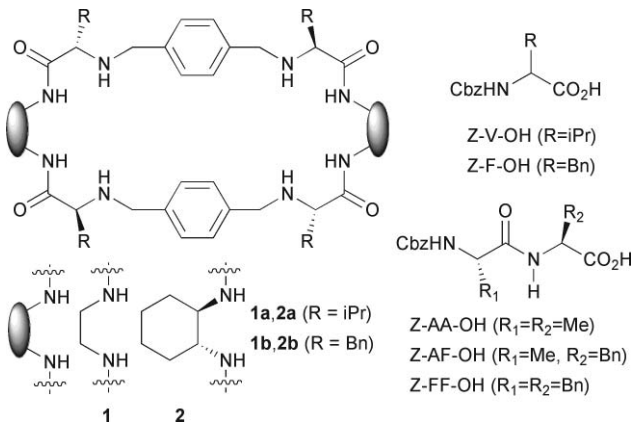

Chart 1

Results and discussion

We initially focused on the use of ESI-MS-based competition experiments as a tool for the high-throughput screening of the solution binding properties of a series of pseudopeptidic macrocycles and some *N*-protected peptide-like molecules, thus leading to a rapid identification of the best receptor for the fittest substrates. We also explored the intrinsic stability of the non-covalent complexes in the gas-phase by tandem mass spectrometry. Following, we have carried out a more in depth structural study of the most interesting examples by NMR, including careful titrations,[31] NOESY[32] and DOSY[33] studies. This comprehensive study has allowed us to compare both techniques and the binding trends in solution and in the gas phase.

Non-covalent interactions screening through competition ESI-MS experiments

As pointed out above, ESI-MS based competition experiments have been widely applied for the extraction of qualitative and/or quantitative information of non-covalent complexes in solution. Essentially, this approach consists of recording the ESI mass spectrum of different host–guest complexes simultaneously. Typically, if a given host (or guest) is mixed with equimolecular amounts of potential guests (or hosts) and the obtained sample is subjected to ESI-MS, the relative intensities of the corresponding supramolecular complexes would give a comparative assessment of the binding differences between host–guest pairs.[34] Besides, if mixed complexes with different host–guest stoichiometries and/or composition are formed, they will be *a priori* detectable. We thus aimed to study the recognition abilities of four different hosts (**1a**, **1b**, **2a** and **2b**) for the binding of Z-protected amino acids (Val and Phe) and dipeptide potential guests.[35] For an initial screening of the macrocycles as receptors, we decided to set up fair competition experiments between the four hosts. Thus, we prepared a stock CH_2Cl_2–CH_3OH (99 : 1) solution of a mixture of equimolecular amounts of the receptors. Small amounts of CH_3OH were required to ensure complete dissolution of the *N*-protected peptides and also proved to slightly enhance the ionization efficiencies (to a similar extent) for the receptors investigated. However, owing to the protic nature of CH_3OH, we used small amounts of methanol (1%) in order to cause minimal disturbance to the potential hydrogen bonding between the hosts and the guests.

Initially, the ESI mass spectrum of the stock solution revealed the presence of singly- ($[M + H]^+$) and doubly-charged ($[M + 2H]^{2+}$) receptors, the latter being dominant. As we shall discuss below, there is no evidence of charged receptors in solution at these conditions as judged by NMR, thus indicating that ionic species are generated during the ESI process. It is remarkable that the branching ratios of the intensity of singly- and doubly-charged species for each receptor were comparable along the series **1a**, **1b**, **2a** and **2b**, thus suggesting identical proton affinities for all of them. Moreover, the peak intensities of the different receptors at parity of concentrations were also comparable. However, this experimental evidence does not guarantee a reliable estimation of solution binding affinities and, accordingly, complementary NMR titration experiments were carried out (see below) to validate the use of ESI-MS based competition methods in the present case.

For competition experiments, the stock solution was mixed with one equivalent of different potential guests in seven separate samples (the final concentration of the guest and every host in all the samples was 10^{-5} M). These samples were directly injected in the spectrometer and analyzed by ESI-MS. Typically all guests investigated formed 1 : 1 complexes with the receptors **1a**, **1b**, **2a** and **2b** as manifested by prominent doubly-charged species of general formula $[R + L + 2H]^{2+}$ together with minor singly charged $[R + L + H]^+$ cations. The obtained results for seven different guests are graphically shown in Fig. 1 taking into consideration the

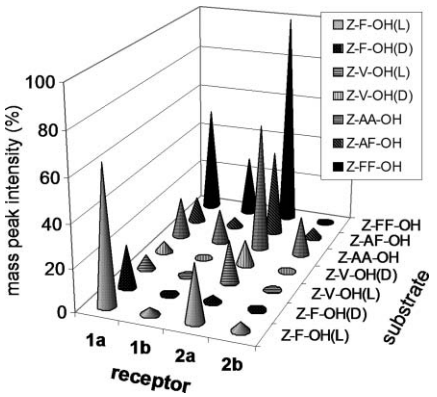

Fig. 1 Plot of the relative peak intensities of the doubly-charged [R + L + 2H]$^{2+}$ species (normalized to 100% of the base peak in the ESI mass spectrum) observed in ESI-MS competition experiments, performed with an equimolecular amount of the four pseudopeptidic receptors and selected peptide-like substrates (all of them 10^{-5} M concentration). Each competition sample is shown with different shading. Note that for the **2a**·*Z*-FF-OH couple, the [**2a** + *Z*-FF-OH]$^{2+}$ dication is the base peak in the ESI mass spectrum, thus anticipating a remarkable binding affinity.

doubly charged non-covalent species. Identical conclusions can be drawn by monitoring the less abundant singly-charged host–guest cations. Complementary ESI mass spectra were recorded with the higher resolution Q-TOF I instrument (operating at a resolution of *ca.* 5000 FWHM) in order to unambiguously ascertain the charge state of the detected species (Fig. S1 of the ESI†).

Two groups of guests were evaluated: (1) *N*-protected amino acids and (2) *N*-protected dipeptides. Regarding the receptors, we observed that the aliphatic side chain is more suitable than the aromatic residues (**1a** > **1b** and **2a** > **2b**). This could be due to an intramolecular interaction of the phenyl ring with the amide/amine/ammonium NH groups of the host, preventing their intermolecular binding with the guest.[36] Among the different substrates, the complexes with the dipeptides seem to be more stable than those with the amino acids, as expected from the macrocyclic ring sizes.[36] Regarding the *N*-protected amino acids, a general trend towards L-selectivity was observed. Additionally, the aromatic Phe also interacts more efficiently with the hosts than the aliphatic Val. On the other hand, the dipeptides showed an increasing binding when additional aromatic groups are added to the molecular structure. Finally, considering all the measured samples, the most versatile and efficient receptor among the tested ones was **2a**, which bears the (*R*,*R*)-cyclohexane-1,2-diyl spacers and *i*-Pr side chains. Thus, these simple competition experiments served us to select receptor **2a** as the most useful structure for additional binding experiments.

To further confirm the performance of **2a** as a receptor, we set up a complementary competition experiment to doubly check the selectivity trends observed in Fig. 1. We prepared an equimolecular mixture of four guests (*Z*-F-OH, *Z*-AA-OH, *Z*-AF-OH and *Z*-FF-OH) and mixed it with the receptor **2a**, achieving the final concentration of 10^{-5} M for host and guests. The mixture was analyzed by ESI-MS and the obtained spectrum is shown in Fig. 2.

Apart from the free receptor, several supramolecular complexes were detected, all of them being both singly- and doubly-charged (produced by the addition of either protons or sodium ions). As described above, we focused our attention on the dominant

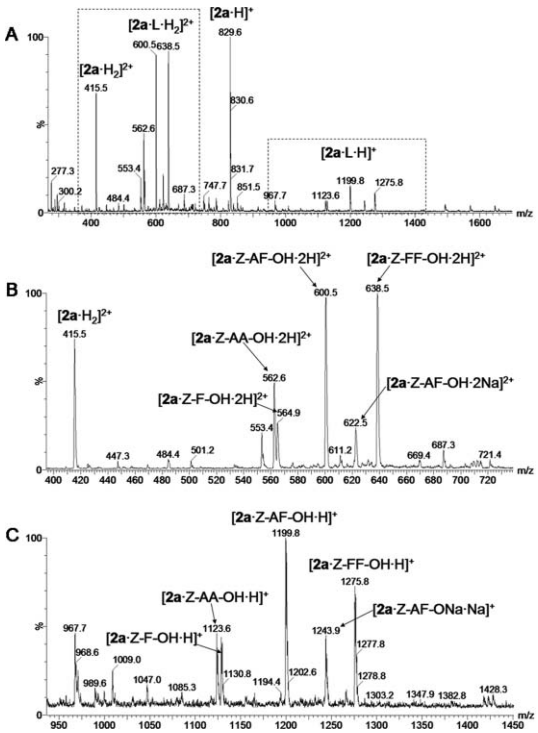

Fig. 2 (a) ESI-MS spectrum obtained by mixing receptor **2a** and an equimolecular mixture of different L substrates (L = *Z*-F-OH + *Z*-AA-OH + *Z*-AF-OH + *Z*-FF-OH), all of them at 10^{-5} M. The dotted regions correspond to the doubly and singly charged supramolecular complex peaks, which are shown enlarged in (b) and (c), respectively. The assignment of some representative peaks is also shown.

[R + L + 2H]$^{2+}$ dications and by comparing the intensities of these peaks, we concluded that the previous observations are also valid in this complementary experiment. In general, dipeptides are more strongly bound to **2a** as the mass peak corresponding to [**2a**·*Z*-F-OH·*n*H]$^{n+}$ (n = 1–2) showed the lowest intensity in the mixture, among all the supramolecular complexes. By comparing the peaks from the complexes with the dipeptides, we also observed that the presence of aromatic side chains increases the strength of the binding (the peaks corresponding to the complexes with *Z*-AF-OH and *Z*-FF-OH are more intense than those for *Z*-AA-OH). Relative abundances of the non-covalent ions were consistent with the binding constants of the host–guest species determined by NMR titration experiments (see Table 1 below). For example, for the receptor **2a**, relative values (normalized to the **2a**·*Z*-FF-OH) derived from peak intensities of the doubly-charged [**2a**·guest + 2H]$^{2+}$ species are **2a**·*Z*-FF-OH (100) > **2a**·*Z*-AF-OH (95) > **2a**·*Z*-AA-OH (50) > **2a**·*Z*-F-OH (25). We also faced the issue of obtaining absolute binding affinities on the basis of ESI-MS measurements following a data treatment method based on 1 : 1 association between the host and guest peptides according the comprehensive work published by Jørgensen *et al.*[37] Our results showed that, although relative binding constants were consistent with NMR titration values, absolute binding constants determined by ESI-MS methods were all found to be systematically higher than the values determined by NMR. Several points may be invoked to account for these differences: among them, distinctive ionization efficiencies of the host and the host–guest complexes, absence of self-aggregation in the ESI-MS measurements (in the

Table 1 Solution association constants (K_{ass}, M^{-1}) obtained from ^1H NMR titration (1% MeOH in $CDCl_3$, 303 K, see ESI† for details) and free energy of interaction (ΔG, kJ mol^{-1}) for the discussed supramolecular complexes

Entry	Receptor	Substrate	K_{ass}/M^{-1}	ΔG/kJ mol^{-1} [a]
1	**2a**	*Z*-AA-OH	86 ± 13	−11.3 ± 0.4
2	**2a**	*Z*-AF-OH	161 ± 19	−12.7 ± 0.3
3	**2a**	*Z*-FF-OH	215 ± 10	−13.4 ± 0.2
4	**2a**	*Z*-F-OH(L)	13 ± 0.5	−6.4 ± 0.1
5	**2a**	*Z*-F-OH(D)	10 ± 0.5	−5.7 ± 0.1
6	**1a**	*Z*-AF-OH	110 ± 9	−11.8 ± 0.3
7	**1a**	*Z*-FF-OH	143 ± 9	−12.4 ± 0.2
8	—[b]	*Z*-AA-OH	53 ± 6	−9.9 ± 0.2
9	—[b]	*Z*-AF-OH	60 ± 5	−10.2 ± 0.2
10	—[b]	*Z*-FF-OH	40 ± 7	−9.2 ± 0.2

[a] $\Delta G = -RT\cdot\ln(K_{ass})$. [b] Corresponding self-aggregation constants.

10^{-5} M range) or the effect of the solvent composition on the formation of the host–guest complex. Overall, considering all the ESI-MS competition experiments, we can extract three main conclusions: (1) the macrocycle **2a** is the best receptor for *N*-protected peptide-like substrates; (2) the complexes are more stable with dipeptides than with simple amino acids and (3) a slight selectivity towards dipeptides bearing aromatic side chains was observed.

Intrinsic gas-phase stability of the dipeptide supramolecular complexes

It is well known that electrostatic interactions are strengthened upon solvent evaporation as the competitive effect of the solvent is removed. Accordingly, it is instructive to probe the intrinsic gas-phase stability of the studied complexes in view to compare them with the information previously gathered in solution.

For this purpose, energy-resolved CID experiments were carried out on ionic complexes formed by **2a** and the *N*-protected dipeptides. For these experiments, the doubly charged complexes, namely $[\mathbf{2a} + L + 2H]^{2+}$ were mass-selected and allowed to collide with an inert gas (argon) at various center-of-mass energies (E_{CM}). In general, the dominant fragmentation channel corresponds to the liberation of the neutral *N*-protected dipeptide yielding the $[\mathbf{2a} + 2H]^{2+}$ doubly-charged free receptor (see eqn (1)). A second minor dissociation pathway is also observed comprising charge splitting to afford the H-adducts $[\mathbf{2a} + H]^+$ and $[L + H]^+$ (see eqn (2)). Product ion spectra recorded at different collision energies are exemplified for the mass-selected $[\mathbf{2a} + Z\text{-FF-OH} + 2H]^{2+}$ dication in Fig. 3 and the rest are given as supporting information (ESI†).

$$[\mathbf{2a} + L + 2H]^{2+} \rightarrow [\mathbf{2a} + 2H]^{2+} + L \quad (1)$$

$$[\mathbf{2a} + L + 2H]^{2+} \rightarrow [\mathbf{2a} + H]^{+} + [L + H]^{+} \quad (2)$$

For a fair and clearer comparison of the data obtained for every complex, the breakdown profiles for each host–guest couple are represented in Fig. 4 by quantifying the decomposition extent of the parent ions in the CID spectrum (see Experimental section and ESI†).

The data clearly show an increase in the complex stability with a larger number of aromatic side chains in the guest. Among the different ways to estimate the intrinsic gas-phase stability of mass-selected species,[38] the $E_{1/2}$ method (collision energy in the center of frame for 50% dissociation of the mass-selected ion) is the most common for dissociating non-covalent complexes.[39] We have estimated the $E_{1/2}$ values for each non-covalent dication as a parameter for comparing relative stabilities. The data clearly indicate the trend *Z*-AA-OH (0.60 eV) < *Z*-AF-OH (0.80 eV) < *Z*-FF-OH (0.85 eV), which suggests that receptor **2a** exerts some selectivity for the dipeptides bearing aromatic side chains. As can be inferred from Fig. 4 and the ESI mass spectrum illustrated in Fig. 2, the results from the CID experiments nicely correlate with those obtained from ESI-MS based competition experiments. For example, from peak intensities in Fig. 2b, it can be estimated that the same order in binding affinities is preserved for the **2a** receptor, thus suggesting a good preservation of the observed selectivity (and thus, the non-covalent interactions) on going from solution to the gas-phase.

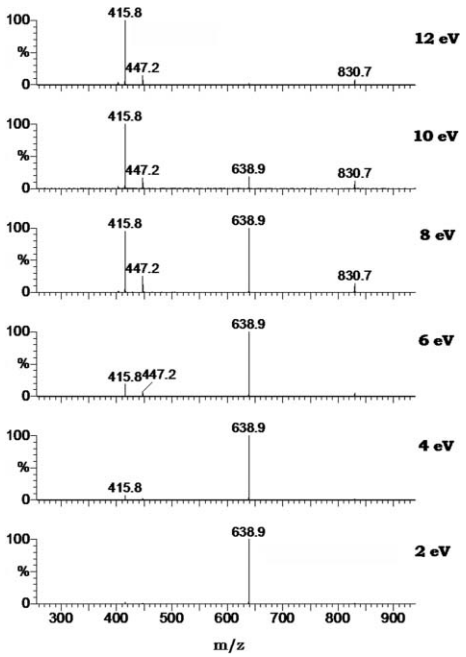

Fig. 3 Collision induced dissociation (CID) spectra of the mass-selected $[\mathbf{2a} + Z\text{-FF-OH} + 2H]^{2+}$ supramolecular complex at increasing collision energies in the $E_{laboratory}$ = 2–12 eV range.

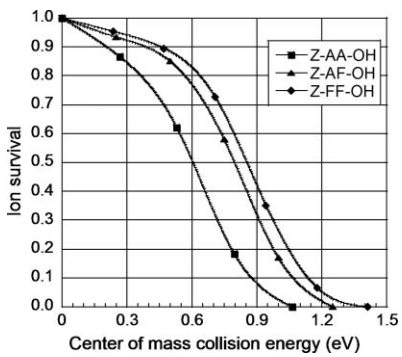

Fig. 4 Breakdown profiles of mass-selected $[\mathbf{2a} + L + 2H]^{2+}$ supramolecular complexes formed by **2a** and either *Z*-AA-OH (squares), *Z*-AF-OH (triangles) or *Z*-FF-OH (diamonds) at increasing center of mass collision energies.

Solution studies by NMR

Considering the results obtained by ESI-MS measurements, we aimed to study the corresponding supramolecular systems in solution by NMR with a double purpose. On the one hand, we aimed to prove the validity of ESI-MS as a high-throughput screening method of the molecular recognition of the investigated pseudopeptidic macrocycles and, on the other hand, we intended to characterize the intimate molecular organization of the non-covalent complexes between the best receptor and selected guests.

NMR titration studies

We initially performed the ^{1}H NMR titration of receptor **2a** ($CDCl_3$, 500 MHz, 303 K) with increasing amounts of *Z*-AF-OH dipeptide. We selected this substrate as it showed the best solubility in organic solvents and a lesser signal overlapping, leading to a clearer difference between the protected amino group and the carboxylic termini. The corresponding ^{1}H NMR spectra for the free host and guest, as well as for a sample of guest-saturated host are shown in Fig. 5.

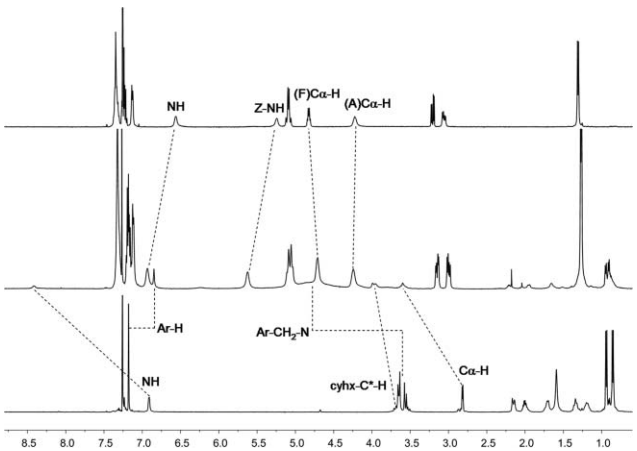

Fig. 5 ^{1}H NMR ($CDCl_3$, 500 MHz, 303 K) spectra of receptor **2a** (lower trace), *Z*-AF-OH dipeptide (upper trace) and a 1 : 15 mixture of **2a**·*Z*-AF-OH (middle trace). The variations of selected signals have been highlighted with dotted lines.

All the corresponding ^{1}H NMR signals were unambiguously assigned by gCOSY, TOCSY and NOESY/ROESY experiments (Fig. S3–S17 of the ESI†). Several proton signals changed their chemical shifts during the titration experiment, which supported the host–guest interactions. The receptor Cα proton signal moves downfield ($\Delta\delta$ = 0.63 ppm) and the benzylic signals (Ar-CH_2–N) changed from an AB quartet centered at 3.6 ppm for the free host to a very broad signal at 4.4–5.2 ppm within the complex. These changes are consistent with a proton transfer from the carboxylic acid of the substrate to the free amino nitrogen atoms of the receptor. Moreover, the proposed acid–base reaction is supported by the slight upfield shift of the Cα proton of the phenylalanine moiety of the dipeptide [(F)Cα-H]. These data suggest the formation of an ionic pair within the supramolecular complex. Other interesting chemical induced shifts are those experienced by the receptor amide NH ($\Delta\delta$ = 1.35 ppm) and cyclohexyl methyne proton signal (cyhx-C*-H, $\Delta\delta$ = 0.27 ppm). These observations are consistent with the participation of the amide N–H groups of **2a** in H-bonding interactions with the substrate, most likely with the dipeptide carboxylic anion. Interestingly, amide, carbamate and alanine Cα protons of the dipeptide also move downfield, suggesting the participation of both amide and carbamate groups in H-bonding with the receptor. Finally, the signal from the aromatic *p*-phenylene proton of **2a** move upfield (Ar–H, $\Delta\delta$ = –0.34 ppm) upon interaction with the substrate, indicating some sort of aromatic–aromatic intermolecular contacts. Also the large broadening of the receptor proton signals upon binding is quite noticeable, suggesting the implication of **2a** in a dynamic process occurring at intermediate rate on the NMR time-scale. The broadening was especially important for the ^{1}H NMR signals of the *p*-phenylene moiety (Ar–H and Ar-CH_2–N), suggesting a slower rotation of this group with respect to the macrocyclic main plane upon the formation of the supramolecular complex. All these data strongly support the formation of a supramolecular complex in solution between **2a** and *Z*-AF-OH, stabilized by electrostatic, H-bonding and aromatic–aromatic non-covalent interactions.

We also intended to get quantitative characterization of the binding using NMR spectroscopy. Unfortunately, the low solubility of the dipeptides (especially *Z*-AA-OH and *Z*-FF-OH) in pure chloroform produced unreliable data due to partial precipitation of the titrant stock solutions. We overcame this problem using 1% methanol as additive. Non-deuterated MeOH was employed in order to allow monitoring NH proton signals, since they were the most suitable for quantitative analysis due to larger chemical induced shift. Besides, since the dipeptides showed some self-aggregation in this medium (see below) we used the dilution method[40] to get quantitative values of binding constants. Under those conditions, the host–guest complexes are favored over the self-aggregation of the corresponding substrates, as during the whole self-titration experiment, the guest has a stoichiometric amount of the host available for the interaction to take place. Moreover, the formed dipeptide carboxylate is expected to have a lesser tendency to self-assemble due to electrostatic repulsion, also favoring the formation of the host–guest complex. We obtained a very good fitting of the NMR titration data to a simple 1 : 1 binding mode (≤15% error with $R^2 > 0.99$, Table 1) in good agreement with the stoichiometry observed by ESI-MS. The observed trends also agreed with the data obtained in the gas phase by ESI tandem MS. Thus, the receptor **2a** binds the dipeptides more tightly than the structurally related amino acid (compare entries 1–3 *versus* 4 in Table 1). According to the solution host–guest binding constants, the interaction is 6–7 kJ mol^{-1} stronger with *N*-protected dipeptides than with a simple *N*-protected amino acid.[17,18,41] Within the dipeptide sequences, the presence of one aromatic side chain additionally increases the stability of the host–guest complex by ~1 kJ mol^{-1} (entries 1–3 in Table 1). Remarkably, by comparing the gas-phase and the solution phase binding interactions we concluded that there is a very good correlation between both states, showing the same stability trends (*Z*-AA-OH < *Z*-AF-OH < *Z*-FF-OH) for the systems studied. Besides, since **2a** is chiral, we studied the stereoselectivity of the molecular recognition process, by performing the NMR titration with the D enantiomer of *Z*-F-OH (entry 5). The data showed a very small L enantioselectivity of the binding (<1 kJ mol^{-1}).

We also performed some NMR titration experiments with macrocycle **1a** bearing a more flexible spacer, which had showed

to be the second best receptor by ESI-MS (Fig. 1). Once again, the NMR titration data nicely correlate with the ESI-MS competition experiments, since **1a** forms slightly less stable complexes with the studied dipeptides (compare entries 2 *vs.* 6 and 3 *vs.* 7 in Table 1). Moreover, the results are strongly consistent, suggesting a stabilization of ~1 kJ mol^{-1} of the host–guest complexes due to the presence of the cyclohexane ring within the receptor structure.

However, a cautionary word must be added for the binding data shown in Table 1. Although we observe moderate binding constants, they have been obtained under conditions where the interactions are weakened by the solvent mixture used. Besides, the self-aggregation of the dipeptides (see below) can partially compete with the host–guest interactions. Therefore, the obtained association constant should be apparent and smaller than the true binding constant.[42]

Self-aggregation of the *N*-protected dipeptides

During the NMR titration experiments, we observed the formation of a fiber-like precipitate in the stock solutions of *Z*-FF-OH. We hypothesized that the observed fibrils might be due to the self-aggregation of the dipeptide. Diphenylalanine dipeptide derivatives are well known as supramolecular self-assembling systems.[43] This peptidic motif has been studied as a model for the formation of fibers related to some neurological disorders, as it is present in the recognition sequence of the pathological aggregates of Alzheimer's disease.[44] Thus, we decided to characterize the self-assembling abilities of the three studied dipeptides under our experimental conditions. Hence, ^{1}H NMR spectra of the dipeptides (1% MeOH in $CDCl_3$, 303 K) at different overall concentrations were investigated. Several signals changed their chemical shifts in this dilution experiments, especially the amide and carbamate NH protons, which suggest the presence of intermolecular NH···O=C hydrogen bonding. We could fit the variation of the amide NH chemical shift to the simplest monomer–dimer equilibrium, rendering an estimation of the corresponding dimerization binding constants (entries 8–10 in Table 1 and ESI†). Very similar dimerization properties were obtained for *Z*-AA-OH and *Z*-AF-OH, but a slightly weaker interaction was observed for *Z*-FF-OH. These differences could be due to the lower solubility of *Z*-FF-OH, which starts to precipitate at concentrations >30 mM. Although we obtained a good fitting of the data, we believe that the simple dimerization model must be an oversimplification of the true situation, since we observed the formation of fiber-like aggregates at higher concentrations.[45] Similar dilution experiments carried out with the receptors showed that **1a** and **2a** do not self-aggregate in a wide concentration range (1–50 mM).

The solid-state structural characterization of a dipeptide would shed some light into the aggregation process observed by NMR. We were able to grow crystals suitable for X ray diffraction analysis, after the very slow evaporation of a solution of the *Z*-AF-OH dipeptide in $CHCl_3$ (Fig. 6). An analysis of the crystal packing of structure *Z*-AF-OH reveals a complex pattern of intermolecular contacts *via* several H-bonding and hydrophobic/aryl–aryl contacts. The carboxylic COOH groups are strongly H-bound (as evidenced by short intermolecular O···O contacts below 2.9 Å), thus forming an infinite chain along the *c* crystallographic axis (Fig. 6A). Within this H-bond

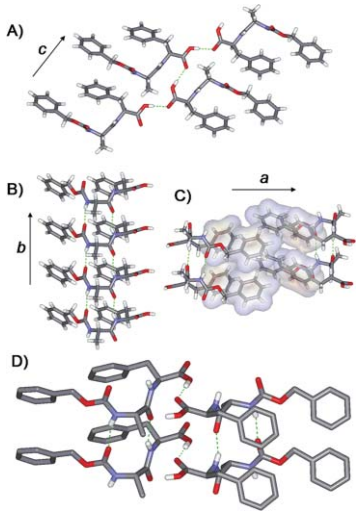

Fig. 6 X-Ray diffraction analysis of the crystal structure of *Z*-AF-OH showing the corresponding interactions along the (A) *c*, (B) *b* or (C) *a* crystallographic axis. (D) Detail of the carboxylic H-bonded dimers of β-sheet-like dimers as the proposed aggregation observed in solution.

pattern, the implicated hydrogen atoms are disordered between the two oxygen atoms of the carboxylic group. Perpendicular to this interaction, the dipeptide molecules are H bound through amide and carbamate N–H···O=C hydrogen bonds (short intermolecular N···O distances are 2.853 Å and 2.925 Å for amide and carbamate groups, respectively) along the *b* crystallographic axis, leading to the formation of an incipient parallel β-sheet-like structure (Fig. 6B). A third level of interactions is established between the aromatic rings of the benzyl side chain of Phe and the *Z*-protecting group. These rings interact through CH–π and π–π contacts (corresponding distances ~3.47–3.87 Å), forming a hydrophobic core which additionally stabilizes the crystal structure along the *a* crystallographic axis (Fig. 6C). Considering the interactions found in the solid state and the ^{1}H NMR signals which mainly changed when acquiring the corresponding spectra at different concentrations, we reasoned that the self-assembling observed in solution is mainly dictated by the amide and carbamate intermolecular H-bonding interactions, leading to β-sheet-like H-bonded dimers. These dimers could further interact through the carboxylic acid proton to yield dimers of dimers (Fig. 6D) which, upon concentration, would eventually lead to the formation of the infinite aggregation observed in the crystal structure of *Z*-AF-OH.

Detailed structural characterization of the supramolecular complex in solution

We also approached the study of the structure of the supramolecular complex in solution with the help of high resolution NMR experiments, which can give detailed additional structural information. To do that, we focused on the [**2a**·*Z*-AF-OH] complex as it showed the best solubility in pure chloroform, avoiding the use of mixtures of solvents. We intended to gather information about two main issues. The first is the binding mode and the host–guest relative disposition. The second concerns the size, shape and dynamics of the host–guest system. In order to study those topics, we have used a series of NOE and diffusion based experiments. Thus, we performed Nuclear Overhauser Effect SpectroscopY

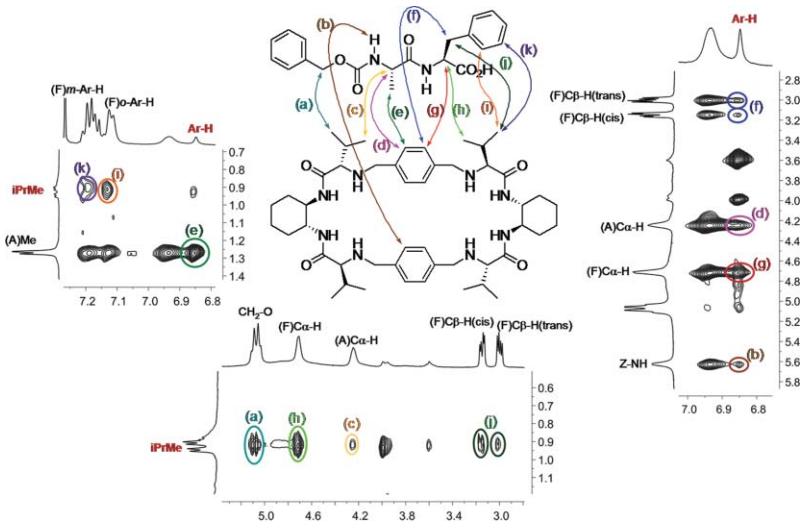

Fig. 7 Selected regions of the 2D NOESY spectrum ($CDCl_3$, 500 MHz, 303 K) of **2a** in the presence of a large excess of *Z*-AF-OH dipeptide (>15 equivalents). The most important intermolecular NOEs are those implicating *p*-phenylene (Ar–H) and isopropyl methyl (*i*-PrMe) protons from the receptor. Key cross-peaks are encircled in colors corresponding to the arrows in the chemical structure, also showing the corresponding lettering.

(NOESY) experiments on several representative samples (Fig. 7 and S13–S17 of the ESI†). The ^{1}H NMR and 2D NOESY data of **2a** show an effective average D_2 symmetry in solution. The coupling constant pattern and NOESY cross-peaks of the cyclohexane moiety imply that it is in a chair conformation, setting the two amide nitrogens in *trans* diequatorial positions. The *i*-Pr side chains would be set in pseudoequatorial disposition regarding the macrocyclic ring. The *p*-phenylene protons appear as a sharp singlet, which suggests a fast rotation of this group with respect to the macrocycle. Interestingly, the 2D NOESY spectrum of a sample of **2a** saturated with *Z*-AF-OH (Fig. 7A, and S15–S17 of the ESI†) showed large negative cross-peaks, while both receptor (Fig. S13–S14) and substrate (Fig. S7) alone showed positive NOEs.[46] This observation supports the formation of large supramolecular species with slow tumbling in solution, and thus, the molecular recognition process is accompanied with a decrease of the rotational movement rate (longer correlation time) of the species.[47] This effect could be due to a combination of an increased effective size, and to the presence of charged species which would produce stronger interactions with the solvent molecules. Selected regions of the 2D NOESY spectrum of this complex are shown in Fig. 7. The observed NOEs in the 2D NOESY experiments were further confirmed by performing the corresponding 2D ROESY measurements (Fig. S16), in order to discard any possible spin diffusion effects.[48] We were able to detect key intermolecular NOEs between the receptor and the substrate. The most diagnostic intermolecular contacts are those implicating the aromatic *p*-phenylene (Ar–H) and side chain (*i*-PrMe) protons of **2a**. The Ar–H protons correlated more strongly with the central part of the dipeptide (both α and β protons of A and F amino acidic residues). On the other hand, the methyl protons of the *i*-Pr side chains are closer to both ends of the linear dipeptide (OCH_2 of the Z protecting group and carboxy terminus). From NOE experiments, we observed a decrease in the rotational rate of the species upon the formation of the supramolecular complex. This experimental evidence prompted us to investigate the changes in the translational movement of the corresponding species. In this sense, the most suitable NMR technique should be the measurement of the corresponding self-diffusion rates by pulse-field gradient spin-echo sequences and, for a clearer comparative representation, their processing using Diffusion Ordered SpectroscopY (DOSY).[49] In these pseudo-2D spectra, the chemical shift is represented in one dimension while the translational self-diffusion rate is shown in the second dimension. Attending to the Stokes–Einstein equation,[50] the translational self-diffusion rate (D) is a physical parameter which depends on the effective molecular dimensions (hydrodynamic radius, r_H), the solvent viscosity (η) and the temperature (T). Therefore, measuring under the same experimental conditions (solvent and temperature) the variations of the D value directly reflects the changes in the effective molecular size. Thus, we initially performed the DOSY experiments (superposed spectra in Fig. 8) with 3.3 mM samples of either **2a** (blue) or *Z*-AF-OH (black), both in $CDCl_3$ at 298 K. In the two cases, the corresponding proton signals of each compound diffuse at a single rate (D), which rendered an average value of $7.6 \pm 0.2 \times 10^{-6}$ cm^2 s^{-1} for *Z*-AF-OH (black in Fig. 8) and $6.0 \pm 0.2 \times 10^{-6}$ cm^2 s^{-1} for **2a** (blue in Fig. 8). These data suggest that the effective molecular size of **2a** is *ca.* double of that for *Z*-AF-OH, in reasonable agreement with the actual molecular sizes of both compounds as monomers. Interestingly, when measuring DOSY experiments with a 30 mM solution of *Z*-AF-OH, we obtained a value of $D = 5.9 \pm 0.1 \times 10^{-6}$ cm^2 s^{-1}, which further supports the self-aggregation of *Z*-AF-OH in more concentrated solutions, as shown by ^{1}H NMR spectra at different concentrations. We then measured a 1 : 1 mixture of **2a** and *Z*-AF-OH (both at 3.3 mM), rendering an even slower translational self-diffusion rate (red spectrum in Fig. 8). Remarkably, the signals corresponding to both (host and guest) diffuse at the same value of D which strongly supports that they are forming a supramolecular complex with a common apparent molecular size. Besides, the obtained value ($D = 4.9 \pm 0.3 \times 10^{-6}$ cm^2 s^{-1}) is smaller than those obtained for both receptor and dipeptide, when isolated at the same concentration and temperature. These data reflect an increase of the effective

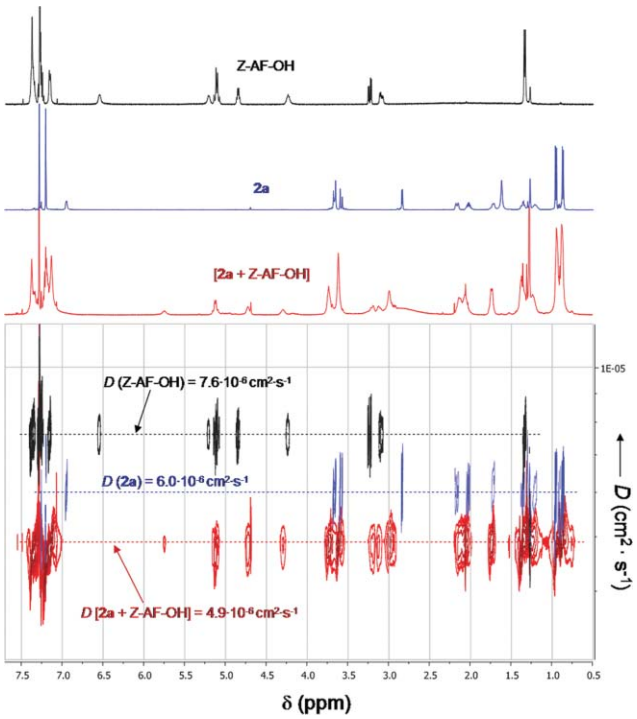

Fig. 8 Superposition of the DOSY and ^{1}H NMR spectra (500 MHz, $CDCl_3$, 298 K) of 3.3 mM solutions of either *Z*-AF-OH (black), **2a** (blue) or a 1 : 1 mixture of *Z*-AF-OH·**2a** (red). The averaged values of the corresponding self-diffusion rates are also given for every sample.

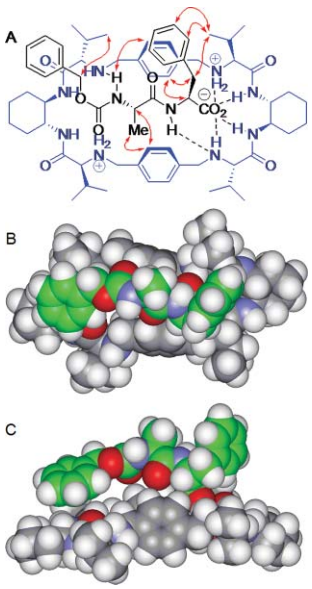

Fig. 9 (a) Schematic representation of the proposed model for the interaction between **2a** and *Z*-AF-OH. Observed intermolecular NOEs are shown by red double-headed arrows. (b) Upper and (c) side views of CPK model of the minimized structure of the same complex. For clarity, the carbon atoms of the dipeptide are shown in green.

molecular sizes of both macrocycle and dipeptide, upon the formation of the supramolecular complex, and again clearly imply the non-covalent host–guest interaction in solution.

Considering the NOEs and the observed chemical induced shifts, we can propose a reasonable model for the supramolecular complex (Fig. 9A), stabilized by electrostatic and H-bonding interactions, with the participation of the aromatic rings of host and guest. The NOE contacts suggest that the dipeptide is located on top of the macrocyclic cavity and aligned with the phenylene groups, setting the central part of the substrate close to these aromatic rings of the host (Fig. 9A).

We used these NOE contacts as constraints for a Monte Carlo conformational search of the complex with MMFF minimizations. The obtained minimum is shown in Fig. 9B–C, which reflects the good host–guest structural complementarity. Moreover, we have used this minimum structure for the calculation of the theoretical self-diffusion rate, by using the HYDRONMR software.[51] We obtained a theoretical value of $D = 4.54 \times 10^{-6}$ cm^2 s^{-1}, which is in good agreement with the experimental one, considering the dynamic nature of the binding process and all the assumed approximations.

Conclusions

In this paper, we have comprehensively studied the molecular recognition properties of a series of pseudopeptidic macrocycles towards *N*-protected amino acids and dipeptides, both in solution (by competition ESI-MS and NMR) and in the gas phase (by CID experiments). We used competition ESI-MS experiments for the screening of the supramolecular complexes formed by four different macrocycles and seven different peptide-like substrates. From these experiments, we found **2a** (bearing cyclohexane-1,2-diyl spacers and *i*-Pr side chains) to be the best receptor, also showing stronger binding with *N*-protected dipeptides over simple *N*-protected amino acids (size selectivity). Among the dipeptides studied, **2a** displayed a slight preference towards aromatic containing dipeptides (sequence or side chain selectivity). The intrinsic stabilities of the [**2a**·dipeptide] supramolecular complexes in the gas phase have been additionally studied by CID experiments, rendering a similar trend to that found by solution competition ESI-MS (*Z*-AA-OH < *Z*-AF-OH < *Z*-FF-OH) and thus, demonstrating the internal consistency of the ESI-MS measurements.

In order to validate the solution ESI-MS data, we additionally performed NMR titration experiments. During this study, we observed the self-aggregation of the dipeptide substrates in solution. The self-assembling of the Z-protected dipeptides occurs through amide and carbamate H-bonding. This proposal has been further supported by the crystal structure of *Z*-AF-OH. Regarding the host–guest complexes stabilities obtained in solution by NMR, for **2a** we observed the same trend as by ESI-MS: *Z*-F-OH ≪ *Z*-AA-OH < *Z*-AF-OH < *Z*-FF-OH. Moreover, we also obtained slightly stronger host–guest complexes with **2a** than with **1a** (bearing ethylene spacers). These results validate both approaches as complementary techniques for their use in the supramolecular chemistry of pseudopeptidic compounds.

We additionally completed the study of the supramolecular complexes in solution by high resolution NMR experiments on the most soluble complex [**2a**·*Z*-AF-OH]. Considering all the NMR data (titrations, NOESY, DOSY) and with the help of molecular modeling, we have been able to propose a model for the host–guest binding.

Experimental

Compound **1a,b**–**2a,b** were synthesized as previously described and showed the expected spectroscopic and analytical data.[30] The Z-protected amino acids and dipeptides are commercially available in pure forms and were used as supplied.

ESI-MS

A Quattro LC (QhQ quadrupole–hexapole–quadrupole) mass spectrometer with an orthogonal *Z*-spray–electrospray interface (Waters, Manchester, UK) was used. Sample solutions approx. 1×10^{-5} M) in dichloromethane: methanol (99 : 1) were introduced through a fused-silica capillary to the ESI source *via* syringe pump at a flow rate of 10 μL min^{-1}. The drying gas as well as nebulizing gas was nitrogen at a flow of 300 L h^{-1} and 80 L h^{-1} respectively. The temperature of the source block was set to 80 °C and the interface to 120 °C. The capillary voltage was set at 3.5 kV in the positive scan mode and the cone voltage was adjusted (typically to a low value $U_c = 15$ V) to control the extent of fragmentation in the source region. The extractor cone and the radio-frequency lens voltage were kept at 3 V and 0.2 V, respectively. Complementary ESI mass spectra were acquired using a Q-TOF I tandem mass spectrometer (quadrupole–hexapole–time-of-flight) operating at a resolution of *ca.* 5000 (FWHM) in order to unambiguously determine the charge state of the ionized species. The chemical composition of each peak obtained in the full scan mode was assigned by comparison of the isotope experimental and theoretical patterns using the MassLynx 4.0 program.

For competition experiments, we used peak areas of species of interest using the triple quadrupole tandem mass spectrometer because of its inherent enhanced dynamic range over TOF analyzers. Collision induced dissociation (CID) experiments were performed using the Q-TOF I tandem mass spectrometer. Typically, the monoisotopic peak of interest was mass-selected with Q1 (isolation width 1 Da), interacted with argon in the hexapole collision cell while analyzing the ionic fragments with the TOF analyzer. The collision energy was systematically stepped in the E_{lab} = 2–14 eV range. For a qualitative analysis of the energy-dependent CID experiments, the laboratory collision energies were converted to the center-of-mass frame, $E_{CM} = m/(m + M)\cdot E_{lab}$, where m and M stand for the masses of the collision gas and the ionic species, respectively. For the breakdown profile representations, signal intensities were obtained from the average of 20 scans and measuring the area of the fragmentation peaks. These graphs were represented taking into account the relative abundance of the precursor and product peaks of each compound ($I_{\text{precursor ion}}/[I_{\text{precursor ion}} + \sum I_{\text{product ion}}]$) against E_{CM}. We selected the value of the collision energy required for 50% reduction of the precursor ion ($E_{1/2}$) as a semi-quantitative measure of intrinsic gas-phase stability of the studied non-covalent complexes.

NMR

All the NMR spectra were performed in a Varian INOVA 500 operating at 500 MHz for proton. Chemical shifts are reported in ppm using TMS as internal standards. The proton signals were unambiguously assigned by means of gCOSY, TOCSY, NOESY and ROESY experiments (these two last ones acquired using 500 ms of mixing time). For the diffusion measurements, the standard DgcsteSL (DOSY Gradient Compensated Stimulated Echo with Spin Lock) sequence was used. The diffusion parameters were optimized for every sample in order to obtain a 90–95% on signal intensity decay. An array of 30 values of gradient strength was acquired with 32–128 scans per value (depending on the sample concentration). The data were processed either with the DOSY macro available in the Varian NMR software or with the MestReNova 5.3.0 software (Mestrelab Research S.L.) rendering essentially the same values. Detailed description of the NMR titration procedures are given in the Electronic Supplementary Information.†

X-Ray crystallographic analysis

Intensity data were collected at 173 K on a STOE IPDS two circle diffractometer with graphite monochromated line MoKα radiation ($\lambda = 0.71073$ Å). In the absence of anomalous scatterers, Friedel pairs were merged. The structure was solved by direct methods (SHELXS)[52] and refined with full-matrix least-squares on F^2 using SHELXL1. All H atoms were located in a difference map. The H atoms bonded to C and O were refined using a riding model. The H atoms bonded to N were freely refined. The hydroxyl H atom is disordered over two equally occupied positions. Crystallographic data: colourless block, $0.19 \times 0.17 \times 0.14$ mm^3, monoclinic crystal system, space group C_2, $a = 43.977(3)$ Å, $b = 5.0960(3)$ Å, $c = 8.3187(6)$ Å, $\beta = 95.848(6)°$, $V = 1854.6(2)$ $Å^3$, $Z = 4$, $\mu = 0.096$ mm^{-1}, 7601 reflections collected, 1833 unique reflections, R_{int} 0 0.0762, θ range 1.86 to 25.0°; R_1 (all data) 0.058, wR_2 (all data) = 0.1607, GooF = 1.151, max. difference peak 0.370, deepest hole –0.219 e $Å^{-3}$.

Molecular modeling

All the modeling studies were performed with Spartan '06 software (Version 1.1.2). Monte Carlo conformational searches were carried out with MMFF force field minimizations and NOE constraints, when applicable. Thus, over 10000 structures were generated and minimized with the MMFF force field. The geometries not satisfying the NOE distances were eliminated from the conformational ensemble. The obtained minima were ordered attending to their relative stabilities and the global minimum was thus located and identified. The corresponding molecular volumes were estimated using the same package.

Acknowledgements

This work was supported by the Spanish Ministerio de Ciencia e Innovación (CTQ2009-14366-C02). M. B. thanks M. E. C. for personal financial support (F. P. U.). The authors are also grateful to the Serveis Centrals d'Instrumentació Científica (SCIC) of the Universitat Jaume I for providing us with mass spectrometric facilities.

Notes and references

1 (*a*) I. Cherny and E. Gazit, *Angew. Chem., Int. Ed.*, 2008, **47**, 4062; (*b*) R. V. Ulijn and A. M. Smith, *Chem. Soc. Rev.*, 2008, **37**, 664; (*c*) E. Gazit, *Chem. Soc. Rev.*, 2007, **36**, 1263; (*d*) C. H. Görbitz, *Chem.–Eur. J.*, 2007, **13**, 1022; (*e*) D. W. Wendell, J. Patti and C. D. Montemagno,

Small, 2006, **2**, 1324; (*f*) X. Zhao and S. Zhang, *Chem. Soc. Rev.*, 2006, **35**, 1105; (*g*) J. M. Lehn, *Science*, 2002, **295**, 2400; (*h*) G. M. Whitesides and B. Grybowsky, *Science*, 2002, **295**, 2418; (*i*) M. D. Hollingsworth, *Science*, 2002, **295**, 2410.

2 (*a*) E. H. C. Bromley, K. Channon, E. Moutevelis and D. N. Woolfson, *ACS Chem. Biol.*, 2008, **3**, 38; (*b*) T. D. Sargeant, M. S. Rao, C. Y. Koh and S. I. Stupp, *Biomaterials*, 2008, **29**, 1085; (*c*) Y. Chau, Y. Luo, A. C. Y. Cheung, Y. Nagai, S. Zhang, J. B. Kobler, S. M. Zeitels and R. Langer, *Biomaterials*, 2008, **29**, 1713; (*d*) K. V. Brinda and S. Vishveshwara, *Biophys. J.*, 2005, **89**, 4159; (*e*) J. R. Heal, G. W. Roberts, J. G. Raynes, A. Bhakoo and A. D. Miller, *ChemBioChem*, 2002, **3**, 136.

3 (*a*) H. Yin and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2005, **44**, 4130; (*b*) Y. Dou, P.-F. Baisnée, G. Pollastri, Y. Pécout, J. Nowick and P. Baldi1, *Bioinformatics*, 2004, **20**, 2767; (*c*) M. W. Peczuh and A. D. Hamilton, *Chem. Rev.*, 2000, **100**, 2479.

4 (*a*) S. Shangary and S. Wang, *Clin. Cancer Res.*, 2008, **14**, 5318; (*b*) K. D. Stigers, M. J. Soth and J. S. Nowick, *Curr. Opin. Chem. Biol.*, 1999, **3**, 714.

5 (*a*) R. S. Root-Bernstein and P. F. Dillon, *Curr. Pharm. Des.*, 2008, **14**, 55; (*b*) H. Yusuf-Makagiansar, M. E. Anderson, T. V. Yakovleva, J. S. Murray and T. J. Siahaan, *Med. Res. Rev.*, 2002, **22**, 146.

6 (*a*) T. Takahashi and H. Mihara, *Acc. Chem. Res.*, 2008, **41**, 1309; (*b*) S. Levin and J. S. Nowick, *J. Am. Chem. Soc.*, 2007, **129**, 13043.

7 D. J. Selkoe, *Cell Biol.*, 2004, **6**, 1054.

8 (*a*) M. F. Perutz, B. J. Pope, D. Owen, E. E. Wanker and E. Sherzinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5596; (*b*) Y. Trottier, Y. Lutz, G. Stevanin, G. Imbert, D. Devys, G. Cancel, F. Saudou, C. Weber, G. David, L. Tora, Y. Agid, A. Brice and J. L. Mandel, *Nature*, 1995, **378**, 403.

9 (*a*) X. Hou, M.-I. Aguilar and D. H. Small, *FEBS J.*, 2007, **274**, 1637; (*b*) C. Haass and D. J. Selkoe, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 101; (*c*) K. A. Jellinger, *J. Neural Transm.*, 2006, **113**, 1603; (*d*) A. I. Bush and R. E. Tanzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7317.

10 R. A. DeFronzo, *Diabetes*, 2009, **58**, 773.

11 (*a*) R. D. Süssmuth, *ChemBioChem*, 2002, **3**, 295; (*b*) K. C. Nicolaou, C. N. C. Boddy, S. Bräse and N. Winssinger, *Angew. Chem., Int. Ed.*, 1999, **38**, 2096; (*c*) D. H. Williams and B. Bardsley, *Angew. Chem., Int. Ed.*, 1999, **38**, 1172; (*d*) C. T. Walsh, S. L. Fisher, I.-S. Park, M. Prahalad and Z. Wu, *Chem. Biol.*, 1996, **3**, 21.

12 (*a*) S. Kuphal, R. Bauer and A. K. Bosserhoff, *Cancer Metastasis Rev.*, 2005, **24**, 195; (*b*) R. O. Hynes, *Cell*, 2002, **110**, 673; (*c*) K. Triantafilou, Y. Takada and M. Triantafilou, *Crit. Rev. Immunol.*, 2001, **21**, 311; (*d*) M. J. Humphries, *Biochem. Soc. Trans.*, 2000, **28**, 311.

13 For some selected examples, see:(*a*) C. Schmuck and P. Wich, *Top. Curr. Chem.*, 2007, **277**, 3; (*b*) B. Botta, I. D'Acquarica, G. Delle Monache, D. Subissati, G. Uccello-Barretta, M. Mastrini, S. Nazzi and M. Speranza, *J. Org. Chem.*, 2007, **72**, 9283; (*c*) C. Schmuck and L. Hernández-Folgado, *Org. Biomol. Chem.*, 2007, **5**, 2390; (*d*) C. Schmuck, *Coord. Chem. Rev.*, 2006, **250**, 3053; (*e*) C. Schmuck and M. Heil, *Chem.–Eur. J.*, 2006, **12**, 1339; (*f*) C. Schmuck and P. Wich, *Angew. Chem., Int. Ed.*, 2006, **45**, 4277; (*g*) M. Wehner, D. Janssen, G. Schäfer and T. Schrader, *Eur. J. Org. Chem.*, 2006, 138; (*h*) L. M. Heitmann, A. B. P. Taylor, J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2006, **128**, 12574; (*i*) M. E. Bush, N. D. Bouley and A. R. Urbach, *J. Am. Chem. Soc.*, 2005, **127**, 14511; (*j*) C. Kahle, R. Deubner, C. Schollmayer, J. Scheiber, K. Baumann and U. Holzgrabe, *Eur. J. Org. Chem.*, 2005, 1578; (*k*) H. Yamamura, M. V. Rekharsky, Y. Ishihara, M. Kawai and Y. Inoue, *J. Am. Chem. Soc.*, 2004, **126**, 14224; (*l*) M. Sirish, V. A. Chertkov and H. J. Schneider, *Chem.–Eur. J.*, 2002, **8**, 1181; (*m*) L. Frish, F. Sansone, A. Casnati, R. Ungaro and Y. Cohen, *J. Org. Chem.*, 2000, **65**, 5026; (*n*) J. Dowden, P. D. Edwards, S. S. Flack and J. D. Kilburn, *Chem.–Eur. J.*, 1999, **5**, 79; (*o*) H. Wennemers, S. S. Yoon and W. C. Still, *J. Org. Chem.*, 1995, **60**, 1108; (*p*) H. J. Schneider, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 848; (*q*) S. S. Yoon and W. C. Still, *J. Am. Chem. Soc.*, 1993, **115**, 823; (*r*) M. Famulok, K.-S. Jeong, G. Deslongchamps and J. Rebek, *Angew. Chem., Int. Ed. Engl.*, 1991, **30**, 858.

14 (*a*) C. Schmuck, D. Rupprecht and W. Wienand, *Chem.–Eur. J.*, 2006, **12**, 9186; (*b*) J. Shepherd, T. Gale, K. B. Jensen and J. D. Kilburn, *Chem.–Eur. J.*, 2006, **12**, 713; (*c*) C. Chamorro, J. W. Hofman and R. M. J. Liskamp, *Tetrahedron*, 2004, **60**, 8691; (*d*) K. A. Ahrendt, J. A. Olsen, M. Wakao, J. Trias and J. A. Ellman, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 1683; (*e*) A. Casnati, F. Sansone and R. Ungaro, *Acc. Chem. Res.*, 2003, **36**, 246; (*f*) G. Chiosis and I. G. Boneca, *Science*, 2001, **293**, 1484; (*g*) R. Xu, G. Greiveldinger, L. E. Marenus, A. Cooper and J. A. Ellman, *J. Am. Chem. Soc.*, 1999, **121**, 4898.

15 (*a*) C. Schmuck, P. Frey and M. Heil, *ChemBioChem*, 2005, **6**, 628; (*b*) P. Rzepecki and T. Schrader, *J. Am. Chem. Soc.*, 2005, **127**, 3016; (*c*) C. Schmuck and M. Heil, *ChemBioChem*, 2003, **4**, 1232.

16 (*a*) C. Schmuck, D. Rupprecht, M. Junkers and T. Schrader, *Chem.–Eur. J.*, 2007, **13**, 6864; (*b*) S. Rensing and T. Schrader, *Org. Lett.*, 2002, **4**, 2161.

17 (*a*) H.-J. Schneider, *Angew. Chem., Int. Ed.*, 2009, **48**, 3924; (*b*) M. Kruppa and B. König, *Chem. Rev.*, 2006, **106**, 3520; (*c*) M. Kruppa, C. Mandl, S. Miltschitzky and B. König, *J. Am. Chem. Soc.*, 2005, **127**, 3362; (*d*) M. A. Hossain and H.-J. Schneider, *J. Am. Chem. Soc.*, 1998, **120**, 11208; (*e*) K. Kano, *J. Phys. Org. Chem.*, 1997, **10**, 286.

18 (*a*) C. A. Schalley, *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH: Weinheim, 2007; (*b*) H.-J. Schneider and A. Yatsimirsky, *Principles and Methods in Supramolecular Chemistry*, Wiley, New York, 2000.

19 (*a*) M. Pons and O. Millet, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2001, **38**, 267; (*b*) F. Hacket, S. Simova and H.-J. Schneider, *J. Phys. Org. Chem.*, 2001, **14**, 159.

20 (*a*) N. M. M. Nibbering, *Mass Spectrom. Rev.*, 2006, **25**, 962; (*b*) M. Speranza, M. Satta, S. Piccirillo, F. Rondino, A. Paladini, A. Giardini, A. Filippi and D. Catone, *Mass Spectrom. Rev.*, 2005, **24**, 588; (*c*) M. Speranza, *Int. J. Mass Spectrom.*, 2004, **232**, 277; (*d*) C. A. Schalley, T. Muller, P. Linnartz, M. Witt, M. Schafer and A. Lutzen, *Chem.–Eur. J.*, 2002, **8**, 3538; (*e*) M. Vincenti and A. Irico, *Int. J. Mass Spectrom.*, 2002, **214**, 23; (*f*) C. A. Schalley, *Mass Spectrom. Rev.*, 2001, **20**, 253; (*g*) C. A. Schalley, *Int. J. Mass Spectrom.*, 2000, **194**, 11; (*h*) C. A. Schalley, J. M. Rivera, T. Martin, J. Santamaria, G. Siuzdak and J. Rebek, *Eur. J. Org. Chem.*, 1999, 1325; (*i*) C. A. Schalley, R. K. Castellano, M. S. Brody, D. M. Rudkevich, G. Siuzdak and J. Rebek, *J. Am. Chem. Soc.*, 1999, **121**, 4568; (*j*) C. A. Schalley, T. Martin, U. Obst and J. Rebek, *J. Am. Chem. Soc.*, 1999, **121**, 2133.

21 M. Schafer, *Angew. Chem., Int. Ed.*, 2003, **42**, 1896.

22 C. A. Schalley, *Mass Spectrom. Rev.*, 2001, **20**, 253.

23 C. B. Lebrilla, *Acc. Chem. Res.*, 2001, **34**, 653.

24 (*a*) T. Mitkina, V. P. Fedin, R. Llusar, I. Sorribes and C. Vicent, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 1863; (*b*) I. Osaka, M. Kondou, N. Selvapalam, S. Samal, K. Kim, M. V. Rekharsky, Y. Inoue and R. Arakawa, *J. Mass Spectrom.*, 2006, **41**, 202; (*c*) H. Zhang, E. S. Paulsen, K. E. Walker, K. E. Krakowiak and D. V. Dearden, *J. Am. Chem. Soc.*, 2003, **125**, 9284.

25 (*a*) E. Ventola, K. Rissanen and P. Vainiotalo, *Chem.–Eur. J.*, 2004, **10**, 6152; (*b*) M. Vincenti and A. Irico, *Int. J. Mass Spectrom.*, 2002, **214**, 23.

26 (*a*) J. S. Brodbelt, *Int. J. Mass Spectrom.*, 2000, **200**, 57; (*b*) C. A. Schalley, *Int. J. Mass Spectrom.*, 2000, **194**, 11; (*c*) J. M. Daniel, S. D. Friess, S. Rajagopalan, S. Wendt and R. Zenobi, *Int. J. Mass Spectrom.*, 2002, **216**, 1.

27 (*a*) B. Baytekin, H. T. Baytekin and C. A. Schalley, *Org. Biomol. Chem.*, 2006, **4**, 2825; (*b*) K. Schug, P. Fryčák, N. M. Maier and W. Lindner, *Anal. Chem.*, 2005, **77**, 3660; (*c*) A. Di Tullio, S. Reale and F. De Angelis, *J. Mass Spectrom.*, 2005, **40**, 845.

28 (*a*) N. K. Beyeh, M. Kogej, A. Ahman, K. Rissanen and C. A. Schalley, *Angew. Chem., Int. Ed.*, 2006, **45**, 5214; (*b*) O. Baudoin, F. Gonnet, M.-P. Teulade-Fichou, J.-P. Vigneron, J.-C. Tabet and J.-M. Lehn, *Chem.–Eur. J.*, 1999, **5**, 2762.

29 For some recent reviews on amino acid containing molecular receptors, see: (*a*) S. Kubik, *Chem. Soc. Rev.*, 2009, **38**, 585; (*b*) I. Alfonso, *Mini-Rev. Org. Chem.*, 2008, **5**, 33; (*c*) S. E. Gibson and C. Lecci, *Angew. Chem., Int. Ed.*, 2006, **45**, 1364.

30 For the synthesis of the macrocycles here studied, see: (*a*) I. Alfonso, M. Bolte, M. Bru, M. I. Burguete and S. V. Luis, *Chem.–Eur. J.*, 2008, **14**, 8879; (*b*) I. Alfonso, M. Bolte, M. Bru, M. I. Burguete, S. V. Luis and J. Rubio, *J. Am. Chem. Soc.*, 2008, **130**, 6137; (*c*) M. Bru, I. Alfonso, M. I. Burguete and S. V. Luis, *Angew. Chem., Int. Ed.*, 2006, **45**, 6155; (*d*) M. Bru, I. Alfonso, M. I. Burguete and S. V. Luis, *Tetrahedron Lett.*, 2005, **46**, 7781.

31 L. Fielding, *Tetrahedron*, 2000, **56**, 6151.

32 (*a*) H. Mo and T. C. Pochapsky, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1997, **30**, 1; (*b*) D. Neuhaus and M. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, 2000; (*c*) A. W. Overhauser, *Phys. Rev.*, 1953, **92**, 411.

33 (*a*) P. S. Pregosin, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **49**, 261; (*b*) P. S. Pregosin, P. G. A. Kumar and I. Fernández, *Chem. Rev.*, 2005, **105**, 2977.

34 (*a*) C. A. Schalley, *Int. J. Mass Spectrom.*, 2000, **194**, 11; (*b*) J. M. Daniel, S. D. Friess, S. Rajagopalan, S. Wendt and R. Zenobi, *Int. J. Mass Spectrom.*, 2002, **216**, 1.

35 I. Alfonso, M. I. Burguete, F. Galindo, S. V. Luis and L. Vigara, *J. Org. Chem.*, 2009, **74**, 6130.

36 (*a*) I. Alfonso, M. I. Burguete, F. Galindo, S. V. Luis and L. Vigara, *J. Org. Chem.*, 2007, **72**, 7947; (*b*) I. Alfonso, M. I. Burguete and S. V. Luis, *J. Org. Chem.*, 2006, **71**, 2242.

37 T. J. D. Jørgensen, P. Roepstorff and A. J. R. Heck, *Anal. Chem.*, 1998, **70**, 4427.

38 M. W. Forbes, D. A. Volmer, G. J. Francis and D. K. Böhme, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 779.

39 W. M. David and J. S. Brodbelt, *J. Am. Soc. Mass Spectrom.*, 2003, **14**, 383.

40 (*a*) L. Peters, R. Fröhlich, A. S. F. Boyd and A. Kraft, *J. Org. Chem.*, 2001, **66**, 3291; (*b*) R. S. Macomber, *J. Chem. Educ.*, 1992, **69**, 375; (*c*) H. J. Schneider, R. Kramer, S. Simova and U. Schneider, *J. Am. Chem. Soc.*, 1988, **110**, 6442.

41 Interestingly, Hunter *et al.* have recently reported an increase of the binding interaction of 6 kJ mol^{-1} per each additional H-bond in cooperative systems. See: C. H. Hunter, N. Ihekwaba, M. C. Misuraca, M. D. Segarra-Maset and S. M. Turega, *Chem. Commun.*, 2009, 3964.

42 B. Hamelin and L. Jullien, *J. Chem. Soc., Faraday Trans.*, 1997, **93**, 2153.

43 (*a*) A. Mahler, M. Reches, M. Rechter, S. Cohen and E. Gazit, *Adv. Mater.*, 2006, **18**, 1365; (*b*) C. H. Görbitz, *Chem. Commun.*, 2006, 2332; (*c*) M. Reches and E. Gazit, *Nano Lett.*, 2004, **4**, 581; (*d*) M. Reches and E. Gazit, *Science*, 2003, **300**, 625.

44 (*a*) E. Gazit, *Prion*, 2007, **1**, 32; (*b*) E. Gazit, *Curr. Med. Chem.*, 2002, **9**, 1725; (*c*) E. Gazit, *FASEB J.*, 2002, **16**, 77.

45 It must be emphasized that monomer-dimer (K_{dim}) and infinite isodesmic self-aggregation (K_E) binding modes cannot be clearlydifferentiated exclusively by titration techniques, as both models can be used to fit the same experimental data rendering $K_E = 2K_{dim}$. See, for instance: Y. Tobe, N. Utsumi, K. Kawabata, A. Nagano, K. Adachi, S. Araki, M. Sonoda, K. Hirose and K. Naemura, *J. Am. Chem. Soc.*, 2002, **124**, 5350.

46 As usually defined, we called "positive" and "negative" NOEs considering the diagonal peaks (2D NOESY) or the irradiated signal (1D NOESY) with negative sign.

47 This fact has been widely observed in biomolecular related systems and is the physical fundament of the so-called transferred NOE experiments, often used to detect small ligand to biomolecule (protein, DNA, *etc*…) interactions. However, this effect has not been so frequently reported in the field of classical host–guest chemistry: see, for instance, reference (*13l*) and also: (*a*) B. Fritzinger, I. Moreels, P. Lommens, R. Koole, Z. Hens and J. C. Martins, *J. Am. Chem. Soc.*, 2009, **131**, 3024; (*b*) J. I. Santos, A. Carvalho de Souza, F. J. Cañada, S. Martín-Santamaría, J. P. Kamerling and J. Jiménez-Barbero, *ChemBioChem*, 2009, **10**, 511.

48 (*a*) A. Bax and D. G. Davis, *J. Magn. Reson.*, 1985, **63**, 207; (*b*) A. A. Bothner-By, R. L. Stephens and J.-M. Lee, *J. Am. Chem. Soc.*, 1984, **106**, 811.

49 (*a*) A. Macchioni, G. Ciancaleoni, C. Zuccaccia and D. Zuccaccia, *Chem. Soc. Rev.*, 2008, **37**, 479; (*b*) Y. Cohen, L. Avram and L. Frish, *Angew. Chem., Int. Ed.*, 2005, **44**, 520.

50 Stokes–Einstein equation for simple spherical particles: $D = (\kappa_b T)/(6\pi\eta r_H)$ where D is the self-diffusion rate, r_H is the hydrodynamic radius, κ_b is the Boltzmann constant and η is the viscosity of the medium at a given temperature (T).

51 J. García de la Torre, M. L. Huertas and B. Carrasco, *J. Magn. Reson.*, 2000, **147**, 138. For a downloadable version: http://leonardo.fcu.um.es/macromol/programs/hydronmr/hydronmr.htm.

52 G. M. Sheldrick, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2008, **64**, 112.