Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Journal Name

ARTICLE

Stereoselective recognition of the Ac-Glu-Tyr-OH dipeptide by pseudopeptidic cages

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Enrico Faggi,[a] Cristian Vicent,[b] Santiago V. Luis,*[,c] and Ignacio Alfonso*[,a]

Pseudopeptidic molecular cages are appealing receptors since they can display different polar and non-polar interaction sites in a modular framework and a controlled disposition. Inspired by previous host-guest knowledge, two pseudopeptidic molecular cages based on serine and threonine (**CySer** and **CyThr**, respectively) were designed and synthesized as hosts for the binding of the four possible stereoisomers of the Ac-Glu-Tyr-OH dipeptide, a target sequence of tyrosine kinases. The careful NMR titration experiments in aqueous acetonitrile allowed the determination of the binding constants and reflected a difference in the stability of the corresponding diastereomeric host-guest complexes. The **CySer** cage proved to be slightly more efficient than the **CyThr** counterpart, although both showed similar stereoselectivity trends: LL > DD ≥ LD > DL. This stereoselective binding was retained in the gas phase, as shown by ESI-MS competition experiments using the enantiomer-labelled method (EL), as well as CID experiments. Thus, the MS-determined discriminations follow the same trends observed by NMR, suggesting that the stereoselectivity observed for these systems must be mainly dictated by the polar host-guest interactions. Despite the stereoselective binding of short peptide sequences in competitive media is a challenging issue in supramolecular chemistry, our results demonstrate the power of pseudopeptidic cages in molecular recognition with foreseen implications in chemical biology.

Introduction

The molecular recognition of short peptide sequences is a challenging topic in supramolecular chemistry due to the number of potential interaction sites and the large conformational flexibility of oligopeptides.[1] This issue is even more difficult when the binding process is performed in highly competitive environments like polar organic solvents or aqueous media.[2] The selective recognition of specific peptidic sequences has been reported using classical hosts such as cucurbituril,[3] calixarenes[4] or cyclodextrins.[5] However, few examples of stereoselective binding can be found in the literature.[6] For obvious reasons, structurally well-defined peptide-like structures are appealing synthetic hosts for the recognition of peptides,[7] since some impressive examples can be found in Nature like the Vancomycin/D-Ala-D-Ala system.[8] Following our studies with bioinspired pseudopeptidic molecules as hosts,[9] we have discovered that the pseudopeptidic cages are excellent candidates to obtain a high selectivity for challenging hosts, like chloride anion in partially aqueous medium.[10] They usually show much more efficient binding than the respective monomacrocyclic receptors.[11] Thus, large pseudopeptidic macrobicycles were prepared by adapting well-established macrocyclization protocols.[12] The modularity of the synthesis allowed access to a large diversity of pseudopeptidic cages with defined rigidity, polarity or potential non-covalent interaction sites.[13] Recently, we have shown that some of these cages are able to bind *N*-protected dipeptides in media of different polarity.[14] For example, sequence specific molecular recognition has been observed for dipeptides bearing an aromatic group at the C-terminus using a variety of different analytical techniques.[14] A particularly well-understood host : guest system is the one formed by **CySer** : Ac-L-Glu-L-Tyr-OH, a dipeptide sequence that is target for Tyr kinases. Careful studies using NMR, ESI-MS and fluorescence spectroscopy, with the help of molecular modelling, gave definitive clues about the stabilization of the complex driven by electrostatic contacts, complementary hydrogen bonding, π-π, hydrophobic and steric interactions. Overall, we used these data to build an interaction model that explains the sequence-selectivity experimentally observed. This model is based on the combination of non-covalent interactions that are configurationally dependent and thus, they could cooperatively work in exerting stereoselective recognition. The stereoselective recognition is by far the most difficult selectivity to achieve in supramolecular chemistry studies.[15] Moreover, this task is even more challenging for flexible linear short peptides and in competitive medium,[16] like aqueous mixtures. Encouraged by the sequence-selectivity

[a] Department of Biological Chemistry and Molecular Modeling, IQAC-CSIC, Jordi Girona, 18-26, E-08034, Barcelona, Spain. E-mail: ignacio.alfonso@iqac.csic.es
[b] Serveis Centrals d'Instrumentació Científica, University Jaume I, Avd. Sos Baynat s/n E-12071 Castellón, Spain.
[c] Department of Inorganic and Organic Chemistry, University Jaume I, Avd. Sos Baynat s/n E-12071 Castellón, Spain. E-mail: luiss@uji.es

† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: Full NMR and ESI-MS spectra of new compounds, NMR titration procedures, data and fitting results, molecular modelling details. See DOI: 10.1039/x0xx00000x

previously reported,[14] here we study the molecular recognition of all the possible stereoisomers of the Ac-Glu-Tyr-OH dipeptide by two pseudopeptidic cages: the previously reported Ser derivative (**CySer**) and the corresponding Thr counterpart (**CyThr**). Besides, for a better understanding of the process, we combined solution studies in competitive media (NMR) with gas-phase assays (ESI-MS). The inspection of all the meaningful streochemical possibilities in the substrate and the use of complementary techniques will allow further refining the interaction model for a better design of a new generation of pseudopeptidic cages with more promising and challenging uses in chemical biology.

Results and discussion

Interaction model for the CySer/Ac-L-Glu-L-Tyr-OH binding

Initially, and based on the results previously obtained for the molecular recognition of the Ac-L-Glu-L-Tyr-OH dipeptide using the cage derived from Serine and bearing a cyclohexane rigid spacer (**CySer**)[14] we elaborated an interactional model (Fig.1) for this host : guest system. Molecular mechanics calculations[17] suggested the main interactions for the host-guest contacts defining three main binding sites (BS). Thus, both carboxylate anions would be complexed by a network of electrostatic and H-bonding contacts comprising two of the pseudopeptidic arms of the host and implicating different amide and amine NH, as well as OH groups as H-bond donors. These H-bonds are in good agreement with the complexation-induced chemical shifts observed in the corresponding NMR titration experiments (see below). Moreover, the Ser OH residues of the host also participate in the carboxylate binding, explaining the higher binding constants observed for the cages bearing this amino acid residue. The Tyr residue of the dipeptide substrate would be encapsulated in the cage cavity setting aryl-aryl contacts with the two tripodal aromatic rings of the receptor and, in this configuration, establishing two additional H-bond interactions with the third pseudopetidic arm of the cage (Tyr BS in Fig.1). Again, these stabilizing interactions implicating the Tyr side chain agree with the stronger binding of dipeptides bearing an aromatic residue at the C-terminus.[14] Besides, the phenol caging could explain the experimentally observed quenching of the Tyr fluorescence by the interaction with the cage.[14] On the other hand, the distance between the Tyr residue and the tripodal aromatic ring of the cage also agrees with the intermolecular ROE observed in the ROESY spectrum of the host-guest system.[14] Finally, the N-terminal acetyl would establish an additional H-bond in the complex. Thus, C-terminus, Glu and Tyr residues interact with the host, suggesting this system a good candidate for assaying stereoselective recognition. This analysis also revealed that the analogous cage based on the Thr amino acid could be of interest in this regard.

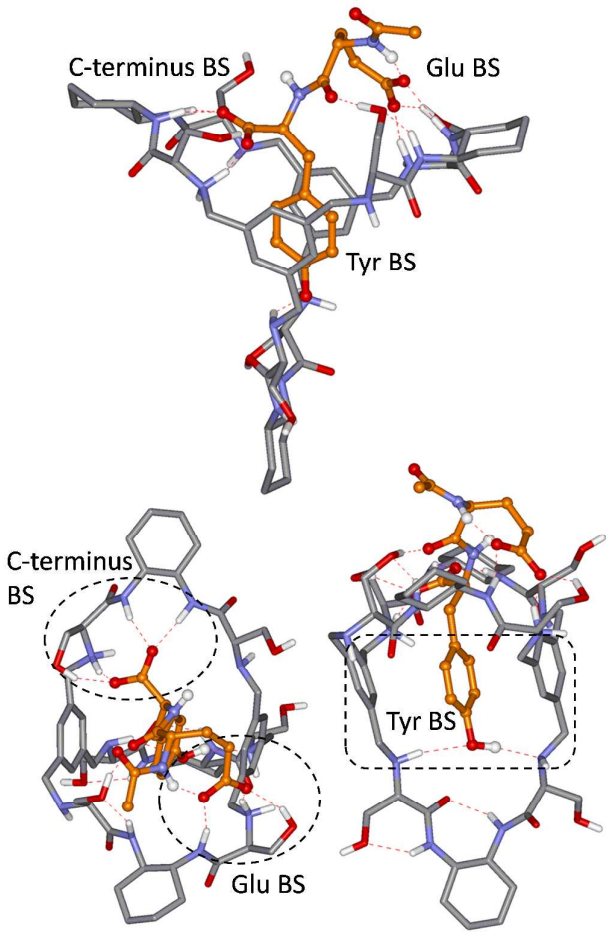


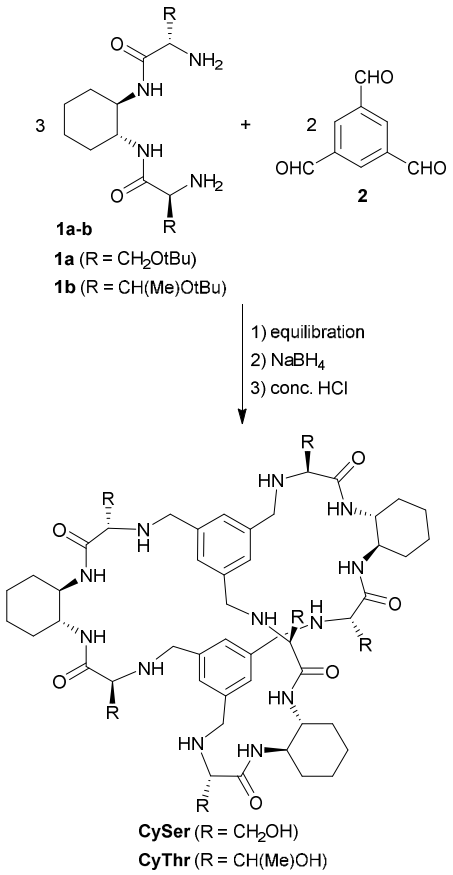
Fig. 1. Binding model for the recognition of Ac-L-Glu-L-Tyr-OH with the **CySer** pseudopeptidic cage. Non-polar hydrogen atoms have been omitted for clarity and the dipeptide substrate is represented with orange C-atoms. The possible H-bonds are represented in red dashed lines and the proposed binding sites (BS) in the host are highlighted with black dashed regions.

Synthesis of the host and guest molecules:

Both pseudopeptidic hosts were synthesized by a reductive amination reaction as shown in Scheme 1, following the procedure already reported for the **CySer** cage.[12,14] The conformationally restricted open-chain bis(amidoamine) **1a,b** derived from Ser and Thr were prepared by conventional amide coupling of (*R*,*R*)-*trans*-cyclohexane-1,2-diamine and the orthogonally protected amino acids (Fmoc-Ser/Thr-*O*tBu), followed by Fmoc deprotection of the α-amino function. Then, the reaction of **1a,b** with benzene-1,3,5-tricarbaldehyde rendered the [3+2] hexaimine cage that was *in situ* submitted to (i) reduction and (ii) deprotection of the side chains, to afford the intended hosts **CySer** and **CyThr** in good overall yields and complete selectivity. The compounds were purified by reverse-phase column chromatography (as their TFA salts) and the free base obtained by reported protocols for the elimination of TFA. The cage identity was supported by the very simple ^{1}H and ^{13}C NMR spectra of **CySer** and **CyThr**, as expected for their high symmetry. Additionally, the

corresponding high-resolution ESI-MS spectra further confirmed the cage structure.

On the other hand, all the dipeptides were synthesized by conventional solid-phase peptide synthesis on Wang resin from the corresponding *N*-Fmoc-protected amino acids, with a final acetylation step for the capping of the *N*-terminus. All the newly synthesized compounds were fully characterized by spectroscopic and analytical techniques (see experimental section).



Scheme 1. Synthesis of the pseudopeptidic cages **CySer** and **CyThr**

Molecular recognition in solution

Nuclear Magnetic Resonance is the most powerful technique for the detailed structural characterization of organic molecules in solution[18] and this is also probably true for the study of supramolecular complexes.[19] Thus, both qualitative and quantitative information can be extracted by carefully planned NMR experiments.[20] From previous screening of a family of cage hosts and different *N*-protected dipeptide guests, we observed that the cage bearing the rigid cyclohexane moiety and derived from serine (**CySer**) was a good receptor for dipeptides having an aromatic amino acid at the C-terminus. Moreover, we observed relatively strong binding toward the Ac-L-Glu-L-Tyr-OH dipeptide, a model sequence that has a biological relevance.[21] This interaction was measured in a medium containing a mixture of polar (acetonitrile) and protic (MeOH) organic solvents. Considering the proposed mode of binding (Fig. 1), we decided to test the stability of this supramolecular complex in a more competitive medium. Thus, we replaced the MeOH component of the initial mixture by water. We reasoned that the hydrophobic component of the supramolecular interaction could be favoured in partially aqueous medium, without an excessive disruption of the electrostatic and polar contacts. Thus, we performed ^{1}H NMR titration experiments of **CySer** with the Ac-L-Glu-L-Tyr-OH dipeptide in CD_3CN containing increasing amounts of water (the use of H_2O allowed monitoring the chemical shifts of the amide protons of host and guest). Several signals from both the host and the guest were perturbed during the titration, supporting that the interaction is also efficient in this very competitive medium (Fig. 2). Amide NH protons from host and guest changed their chemical shift suggesting their participation in intermolecular H-bonds. Also the Ar-H signal from the host is affected by the dipeptide, which implies the participation of the aromatic rings in the recognition. By comparing the titration curves at different water contents, several interesting trends can be observed (Fig.2). For the host, the behaviour of the proton signals is very similar at 33% and 50% of water, both in the absolute values of the chemical shifts and in their variations when adding the guest. This remarkable result suggests that the microscopic environment of the host during the titration experiments is very similar at 33% and 50% of water. In the case of the guest, the different water contents produced larger differences. We reasoned that the linear and flexible nature of the dipeptide leaved the structure more exposed to the solvent and thus, more sensitive to the solvent composition. However, the closed structure of the pseudopeptidic cage seems to have a limiting solvation shell,[22] which is reached at 30% of water. The shape of the titration isotherms suggested the formation of a 1 : 1 host : guest complex, which was further confirmed by the fitting of the data and by ESI-MS experiments.

The simultaneous fitting of the NMR chemical shift variations of the host and the guest rendered association constants in the range of 630-710 M^{-1} (Fig. 3). Interestingly, the stability of the interaction remains practically unaffected by the amount of water present in the solvent, probably due to the dual nature of the host-guest interactions. Another plausible explanation is that the microenvironment of the receptor is very similar regardless the overall amount of water present in the experiment, since this host would show a very efficient hydration sphere.[22-23] This is also in agreement with the plots shown in Fig. 2A,B, where the variations of chemical shifts of the host are very similar at different water contents (actually, almost identical at 33 and 50% water). From the plot shown in Fig. 3, we concluded that the slightly most competitive mixture (lower K_{as}) must be 2 : 1 CD_3CN : H_2O and thus, we decided to carry out all the subsequent studies in this medium.

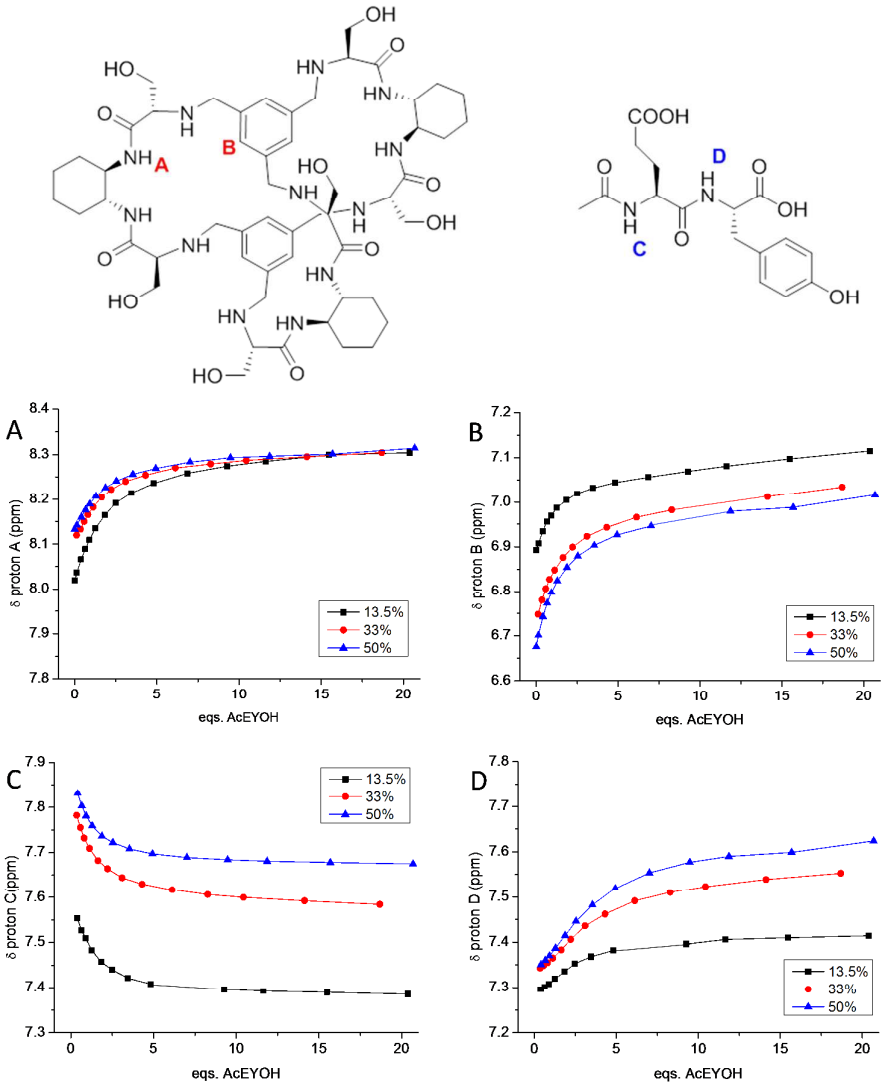


Fig. 2. Plot of the variation of the ^{1}H NMR chemical shifts during the titration of **CySer** cage with Ac-L-Glu-L-Tyr-OH dipeptide (500 MHz, 298 K) at different proportions of water (13.5% black, 33% red or 50% blue) in CD_3CN: (A) host amide proton, (B) host aromatic signal, (C) Glu amide proton from the guest and (D) Tyr amide proton from the guest. The corresponding NMR signals (see structures on the top) were assigned by 2D NMR experiments.

In all the previous studies, we always used dipeptides containing the naturally occurring L-amino acids. Since the **CySer** cage is a chiral host and the host-guest contacts proposed in our model are stereochemically dependent, we envisioned studying the binding of this cage toward all the possible stereoisomers of the dipeptide. Besides, we also assayed the cage derived from threonine (**CyThr**), which also contains the OH groups for H-bonding and adds additional stereogenic centres on the side chains. The simultaneous fitting of the chemical shifts variations for several signals of both hosts and guests, observed in the NMR titration experiments, showed the formation of 1 : 1 supramolecular complexes in all the cases. The host-guest binding constants for the two cages and the four stereoisomers of the dipeptide are shown in Table 1, also with the corresponding Gibbs energy for the process. For a better comparison of the K_{as}, the values are plotted in Fig. 4.

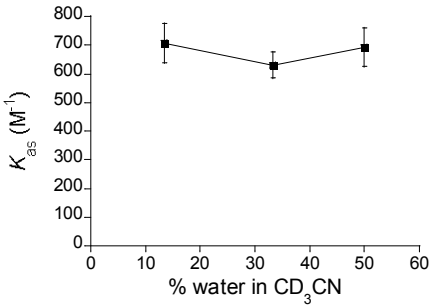


Fig. 3 Plot of the stability constant of the [CySer · Ac-Glu-Tyr-OH] complex (^{1}H NMR titrations, 500 MHz, 298 K) against the % water in CD_3CN.

Table 1. Association constants (K_{as}, M^{-1}) for the interaction between the two pseudopeptidic cages and the four possible stereoisomers of the Ac-Glu-Tyr-OH dipeptide, determined by ^1H NMR (500 MHz, 298 K) titration experiments in 3 : 1 CD_3CN : H_2O solvent.

Cage host	Dipeptide guest	K_{as} (M^{-1})	ΔG (kJ/mol)
CySer	Ac-L-Glu-L-Tyr-OH	631±45	-16.0±0.2
CySer	Ac-D-Glu-D-Tyr-OH	457±33	-15.2±0.2
CySer	Ac-L-Glu-D-Tyr-OH	447±43	-15.1±0.2
CySer	Ac-D-Glu-L-Tyr-OH	282±13	-14.0±0.1
CyThr	Ac-L-Glu-L-Tyr-OH	550±25	-15.6±0.1
CyThr	Ac-D-Glu-D-Tyr-OH	457±22	-15.18±0.05
CyThr	Ac-L-Glu-D-Tyr-OH	339±8	-14.44±0.06
CyThr	Ac-D-Glu-L-Tyr-OH	200±9	-13.1±0.1

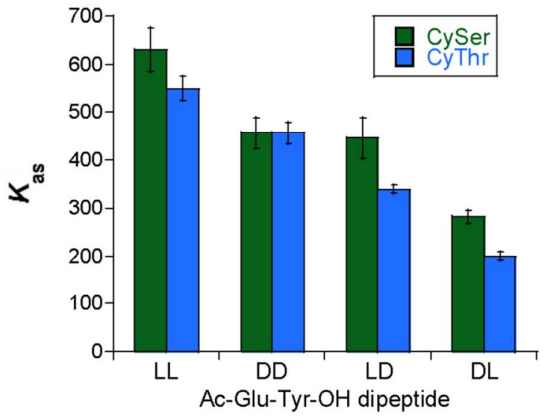


Fig. 4. Plot of the association constant (M^{-1}) for the binding of the four stereoisomers of Ac-Glu-Tyr-OH dipeptide with **CySer** (green) and **CyThr** (blue) hosts, obtained by NMR titrations (33% water in CD_3CN, 500 MHz, 298.15 K).

Despite the moderate stereoselective recognition, most of the differences are above the experimental errors. Accordingly, several meaningful conclusions can be extracted from the data reported in Table 1. First of all, both cages showed stereoselective recognition of the guest dipeptides with a consistent selectivity trend: LL > DD ≥ LD > DL. This fact suggests that the additional stereogenic centres in the side chains of the **CyThr** receptor are not important for the selectivity. On the contrary, since the binding constants with the **CyThr** receptor are generally lower, the substitution in the alcohol side chain would disfavour the host-guest interactions by steric hindrance. The naturally occurring LL isomer of the dipeptide was the most tightly bound by the two cages (also constructed with L amino acids). Interestingly, there is a *match*/*mismatch* effect of the chiral centres of the substrate for the binding to the receptors. Thus, the inversion of the configuration of the Glu residue has a higher impact on the binding when the configuration of the Tyr is L, but a much lower effect (**CyThr**) or even no effect (**CySer**) when the configuration of the Tyr is D. These results imply that the configurations of the two residues of the dipeptide are cooperatively modulating the binding and therefore, both side chains of the substrate must be interacting with the cage, as proposed in our initial model. The most efficient combination in terms of stereoselectivity is the LL/DL pair. Thus, it seems that the H-bonding interactions (mainly established by the Glu side chain) should be critical for the selectivity when the hydrophobic contacts (mainly implicating the Tyr residue) remain unaffected.

Molecular recognition in the gas phase

Mass spectrometric techniques have been extensively used to investigate host-guest molecular recognition processes[24] that also include chiral discrimination. Numerous quantitative enantiomer assays have been demonstrated using soft ionization methods.[25] Upon transferring a supramolecular complex from the solution to the gas phase, enthalpically favoured electrostatic forces become stabilized whereas hydrophobic forces become destabilized.

In the present study, we envisioned instructive to extend the solution molecular recognition studies to the gas phase using ESI-MS because it allows to compare and to better understand the role of solvent and its effect on the hydrophobic, hydrogen-bonding or electrostatic interactions, which modulate the stereoselectivity of the host-guest process. In particular, the enantiomer-labelled (EL) method was chosen to trace chiral recognition by means of ESI-MS.[26] In the EL method, the mass-labelling of one chiral guest and its complexation with the chiral host in competition with the unlabelled guest allows mass-differentiated (pseudo)stereoisomeric complexes to be observed in the ESI mass spectrum where distinctive ion abundances should be indicative of chiral discrimination. Following this approach, we determined the stereoselectivity trends and configurational preferences for association of **CySer** and **CyThr** with the isomers of Ac-Glu-Tyr-OH substrate. Mass-labelling of the Ac-L-Glu-L-Tyr-OH dipeptide (see Experimental section for details) was carried out on the methyl groups of the acetyl moiety incorporated into the *N*-terminal blocking group on the dipeptide.

In order to evaluate the extent of the chiral recognition, equimolecular mixtures of the labelled d^3-Ac-L-Glu-L-Tyr-OH and an unlabelled guest stereoisomer were prepared, mixed with **CySer** (or **CyThr)** and subjected to ESI-MS analysis. The corresponding peaks for the diprotonated host-guest complexes, assigned to [**CySer** + Ac-Glu-Tyr-OH +2H]$^{2+}$ (m/z 723.4) and [**CySer** + d^3-L-Ac-L-Glu-Tyr-OH +2H]$^{2+}$ (m/z 724.9) or [**CyThr** + Ac-Glu-Tyr-OH +2H]$^{2+}$ (m/z 765.4) and [**CyThr** + d^3-L-Ac-L-Glu-Tyr-OH +2H]$^{2+}$ (m/z 766.9) were identified in their respective ESI mass spectra. By the inspection of the corresponding normalized intensities of the ESI mass spectra containing the d^3-Ac-L-Glu-L-Tyr-OH (as the reference) and Ac-D/L-Glu-D/L-Tyr-OH in competition (see Table 2 and Fig. S47), the ion abundances of the three diastereomeric [H:G] species follow the trend: LL > DD ≥ LD > DL for both **CyThr** and **CySer** receptors. The difference is especially large for the binding of the DL/LL pair, suggesting that the configuration of the chiral centre of the Glu residue is more important for the stereoselectivity.

Table 2. Ion abundance ratio corresponding to the diastereomeric complexes incorporating each of the Ac-Glu-Tyr-OH enantiomers and the deuterated d^3-Ac-L-Glu-L-Tyr-OH enantiomer with the host (**CySer** or **CyThr**) of interest.

Guest	CySer[a]	CyThr[b]
Ac-L-Glu-L-Tyr-OH	1.00	1.00
Ac-D-Glu-D-Tyr-OH	0.92	0.81
Ac-L-Glu-D-Tyr-OH	0.84	0.74
Ac-D-Glu-L-Tyr-OH	0.62	0.72

[a]Ion abundance ratio of the [**CySer** + Ac-Glu-Tyr-OH +2H]$^{2+}$ (m/z 723.4) / [**CySer** + d^3-L-Ac-L-Glu-Tyr-OH +2H]$^{2+}$ (m/z 724.9) complexes. [b]Ion abundance ratio of the [**CyThr** + Ac-Glu-Tyr-OH +2H]$^{2+}$ (m/z 765.4) / [**CyThr** + d^3-L-Ac-L-Glu-Tyr-OH +2H]$^{2+}$ (m/z 766.9) complexes. In both cases, the ion abundance ratios are normalized to that of the Ac-L-Glu-L-Tyr-OH enantiomer displaying the highest affinity.

The observed stereoselectivity was also manifested in the CID experiments of mass-selected supramolecular adducts (Fig. 5). For example, the adduct of the **CySer** host with the Ac-L-Glu-L-Tyr guest, namely [CySer + Ac-L-Glu-L-Tyr +2H]$^{2+}$ display enhanced kinetic stability with respect to its stereoisomeric counterparts. More specifically, CID mass spectra of mass-selected adducts [CySer + Ac-Glu-Tyr-OH +2H]$^{2+}$ at $Ce_{Elaboratory}$ = 5 eV induced the following dissociation percent for the different complexes: 35% (LL), 45% (DD), 54% (LD) and 67% (DL), as shown in Fig. 5. Very similar results were obtained with the **CyThr** host (See Fig. S48).

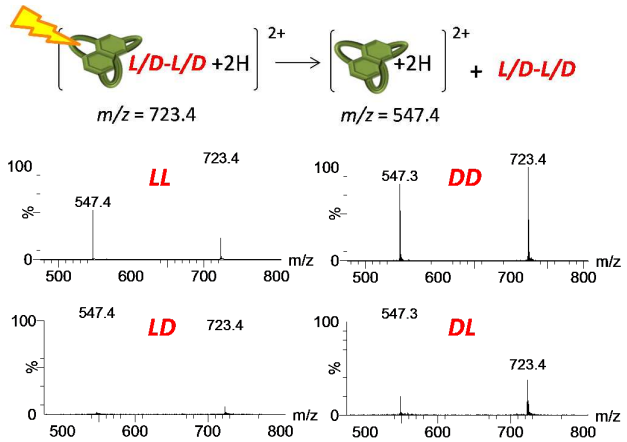


Fig 5. CID mass spectra of mass-selected adducts [**CySer** + Ac-Glu-Tyr-OH +2H]$^{2+}$ at $Ce_{Elaboratory}$ = 5 eV for the corresponding complexes formed by the **CySer** host and the different stereoisomers of the dipeptide.

Overall, ESI-MS analysis showed chiral discrimination of the **CySer** and **CyThr** hosts for the stereoisomeric Ac-Glu-Tyr-OH guests, with a clear preference for the LL dipeptide. This pattern is in excellent agreement with the trends observed with the other techniques. It is reasonable to assume that in this example, if hydrophobic forces due to Tyr encapsulation played the key role at differentiating stereoisomers, ESI-MS analysis would be unable to distinguish between the enantiomeric guests because the hydrophobic forces would be cancelled. Therefore, the trends observed in ESI-MS suggest that the stereoselective recognition is mainly due to polar (electrostatic or H-bonding) interactions, which are retained and even strengthened in the gas phase.

Molecular modelling

Following the proposed structure for the [**CySer** · Ac-L-Glu-L-Tyr-OH] species, we also modelled the corresponding complexes for all the dipeptide stereoisomers (Fig. 6). The molecular mechanics calculations suggest the formation of structurally similar supramolecular complexes, although the different configurations of the substrates affect the non-covalent interactions established with the receptor. These are reflected in the subtle differences for the distances between the different moieties of the guests within the defined BS of the cage (see Fig. S49 and its corresponding caption in the ESI for details). Thus, the observed stereoselectivity trend is a result of the combination of all the non-covalent interactions, where the polar interactions (H-bonds) play major role.

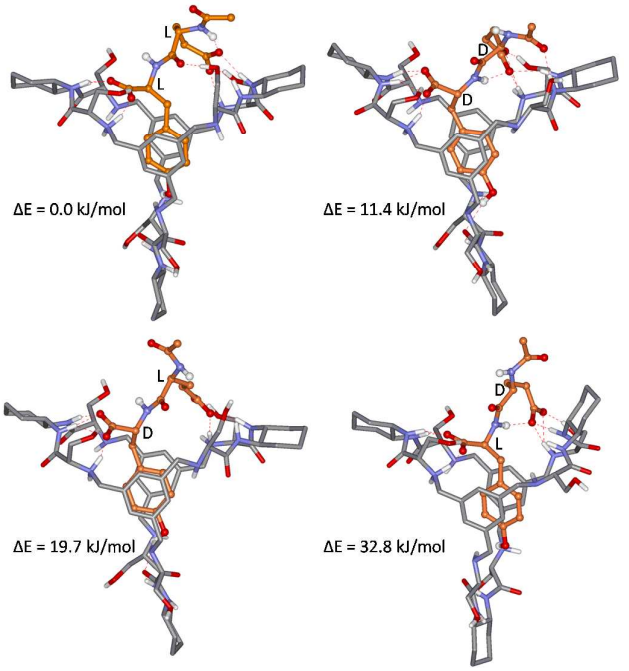


Fig. 6. Molecular mechanics structures for the four diastereomeric complexes formed between the **CySer** cage and the corresponding D/L-D/L dipeptides. Non-polar hydrogen atoms have been omitted for clarity and the dipeptide substrates are represented with orange C-atoms. The possible H-bonds are drawn in red dashed lines and the calculated relative MMFF energies are also included.

The alignment of the minimized geometries also showed interesting features (Fig. 7). Regarding the disposition of the guests, the C-termini and the Tyr residues nicely overlay for the four complexes, while a more evident geometrical difference was obtained for the corresponding *N*-termini and the Glu residues (Ac-Glu in Fig 8). Regarding the host, the complexation with the DL dipeptide produced a larger distortion of the host geometry as compared with the other three complexes. Overall, these structural features could explain the observed stereoselectivity. The geometrical differences are also reflected in the calculated MMFF energies (Fig. 6), which show the same trend as the experimental binding energy. A very similar result was obtained with the **CyThr** cage (See Fig. S50 in the ESI). However, the theoretical

calculations yielded a much larger quantitative difference than the one observed experimentally, which can be due to the overestimation of the binding forces or, most likely, due to the co-existence of less selective binding modes. This last possibility cannot be overruled considering the size and flexibility of our host-guest systems. Thus, although the proposed model is an obvious simplification of the real situation, we found remarkable that it nicely explains the experimental results obtained using two different techniques and in different media.

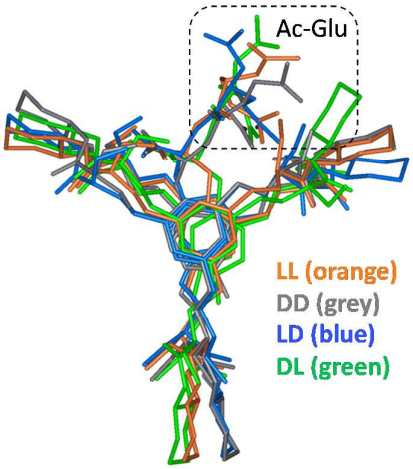


Fig. 7. Aligned geometries for the complexes formed by the **CySer** host and the four diastereomeric dipeptides: LL (orange), DD (grey), LD (blue) and DL (green).

Conclusions

In summary, the stereoselective binding of the Ac-Glu-Tyr-OH dipeptide by two pseudopeptidic cages derived from Ser and Thr, respectively, was studied. The NMR titration studies in a competitive aqueous medium showed a trend in the recognition of the different stereoisomers of the dipeptide (LL > DD ≥ LD > DL), with a clear preference for the binding of the naturally occurring isomer (LL). This stereoselection was retained in the gas phase as confirmed by competition ESI-MS experiments (using an isotopically labelled isomer as standard) and CID assays of the observed ions for the 1 : 1 host : guest complexes. The experimental results, in combination with molecular modelling, allowed proposing a reasonable binding mode for the stereoselective binding, based on different polar and non-polar host-guest interactions. The NMR data at different solvent compositions suggest that both polar and non-polar interactions are important for the stability of the complexes. However, the preservation of the selectivity in the gas phase highlights the key role of the polar H-bonds in the stereoselection. Overall, the combination of solution and gas-phase binding studies has allowed the refining of the model for the interaction, which will lead to the further optimization of synthetic pseudopeptidic cages for targeting more challenging peptide substrates. Work along this line is in progress in our group.

Experimental section

General features

Reagents and solvents were purchased from commercial suppliers (Aldrich, Fluka, or Merck) and were used without further purification. Preparative reverse phase purifications were performed on a BioTage instrument (KP-C18-HS, CH_3CN and water with 0.1% TFA). Analytical RP-HPLC was performed with a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reverse-phase Kromasil 100 C8 (15 x 0.46 cm, 5 µm) column. CH_3CN-H_2O Mixtures containing 0.1% TFA at 1 mL/min were used as mobile phase and monitoring wavelength was set at 220 and 254 nm. The NMR spectroscopic experiments were carried out on a Varian INOVA spectrometers (500 and 400 MHz for ^1H and 125 and 100 MHz for ^{13}C NMR). For the characterization of the new compounds, high resolution mass spectra (HRMS) were performed on Acquity UPLC System and a LCT PremierTM XE Benchtop orthogonal acceleration time-of-flight (oa-TOF) (Waters Corporation, Milford, MA) equipped with an electrospray ionization source.

Synthetic procedures

CySer was synthesized as already reported.[12,14]

$(Fmoc)_2$-1b: Fmoc-Thr(*t*Bu)-OH (3.73 g, 9.40 mmol) was dissolved in dry DMF (10 ml) and dry DCM (10 ml). 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronoium hexafluorphosphate (HATU, 3.41 g, 8.97 mmol) and *N*,*N*-Diisopropylethylamine (DIPEA, 6.0 ml, 34.2 mmol) were added over the solution. The reaction mixture was cooled to 0ºC. A solution of the dihydrochloride salt of (*R,R*)-1,2-diaminocyclohexane (800 mg, 4.27 mmol) in dry DMF (10 ml) was added over the mixture. The solution was allowed to warm to room temperature for 16 hours, after which complete conversion of starting material was observed by TLC. The mixture was diluted with water and extracted with AcOEt (3 x 50 ml). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over $MgSO_4$ and concentrated to dryness. The residue was purified by flash chromatography using hexane: AcOEt as eluent (from 20% to 50% AcOEt) to give 3.10 g of **$(Fmoc)_2$-1b** (83% yield) as a white solid.

^1H NMR (400 MHz, $CDCl_3$) δ 7.74 (d, *J* = 8.3 Hz, 4H), 7.60 (dd, *J* = 7.7, 3.9 Hz, 4H), 7.38 (td, *J* = 7.3, 2.9 Hz, 4H), 7.28 (t, *J* = 7.5 Hz, 4H), 7.22 (d, *J* = 5.5 Hz, 2H), 6.03 (d, *J* = 5.3 Hz, 2H), 4.40-4.31 (m, 4H), 4.23 (t, *J* = 7.3 Hz, 2H), 4.15 (dd, *J* = 6.4, 3.9 Hz, 2H), 4.09 (t, *J* = 4.6 Hz, 2H), 3.68 (s, 2H), 2.09 (d, *J* = 12.7 Hz, 2H), 1.76 (d, *J* = 8.1 Hz, 2H), 1.39-1.34 (m, 2H), 1.28 (s, 18H), 1.02 (d, *J* = 6.3 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.5, 156.0, 144.1, 143.9, 141.4, 127.8, 127.1, 125.3, 120.1, 75.5, 67.1, 66.5, 58.5, 53.6, 47.3, 32.2, 28.2, 24.5, 17.1. HRMS (ESI-TOF) *m/z* [**$(Fmoc)_2$-1b** + H]$^+$ Calcd for $C_{52}H_{65}N_4O_8$ 873.4802, found 873.4832.

1b: $(Fmoc)_2$-1b (3.06 g, 3.51 mmol) was dissolved in 10 mL of 20% Piperidine in DMF. After several minutes the product precipitates as a white solid but the mixture was allowed to react for 6h until complete conversion of starting material was observed by TLC. Excess diethyl ether was added over the reaction mixture and the product was filtered off and washed

with additional diethyl ether. **1b** was obtained as a white solid (1.13 g, 75% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 3.81 (qd, *J* = 6.2, 4.7 Hz, 2H), 3.65-3.60 (m, 2H), 3.13 (d, *J* = 4.8 Hz, 2H), 1.99-1.94 (m, 2H), 1.79-1.74 (m, 2H), 1.38-1.31 (m, 4H), 1.22 (s, 18H), 1.09 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 175.1, 75.4, 70.4, 61.0, 54.4, 33.1, 28.8, 25.7, 19.3. HRMS (ESI-TOF) *m/z* [**1b** + H]$^+$ Calcd for $C_{22}H_{45}N_4O_4$ 429.3441, found 429.3503.

CyThr: A solution of benzene-1,3,5-tricarbaldehyde (116 mg, 0.72 mmol) in CH_3OH (25 ml) was added over a solution of **1b** (461 mg, 1.07 mmol) in CH_3OH (10 ml). The mixture was stirred at room temperature during 20 hours. Then, $NaBH_4$ (163 mg, 4.32 mmol) was carefully added and the mixture was allowed to react for 16 hours. The mixture was concentrated to half volume and *ca.* 2 ml of concentrated HCl was added. After 2 hours the mixture was evaporated to dryness. The residue obtained was dissolved in water and basified with 1N NaOH, the product was extracted with DCM. The combined organic layers were dried over $MgSO_4$ and the solvents were evaporated in vacuum. The product was purified using reversed-phase flash chromatography (eluant: 1% to 20% MeCN in water; 0.1% TFA in both solvents). The TFA salt was transformed into the free-base amine using an ion-exchange resin, affording **CyThr** as a white solid (233 mg, 55% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 6.94 (s, 6H), 3.97 , 3.94 3.94 (p, *J* = 6.3 Hz, 6H), 3.82-3.80 (m, 6H), 3.65 (d, *J* = 12.9 Hz, 6H), 3.42 (d, *J* = 12.9 Hz, 6H), 3.04 (d, *J* = 6.3 Hz, 6H), 2.11-2.09 (m, 6H), 1.84- 1.72 (m, 6H), 1.40-1.32 (m, 12H), 1.23 (d, *J* = 6.3 Hz, 24H). ^{13}C NMR (101 MHz, CD_3OD) δ 174.4, 140.2, 127.4, 69.4, 69.2, 54.0, 53.5, 33.5, 25.8, 20.2. HRMS (ESI-TOF) *m/z* [**CyThr** + H]$^+$ Calcd for $C_{60}H_{97}N_{12}O_{12}$ 1177.7349, found 1177.7354; [**CyThr** + Na]$^+$ Calcd for ($C_{60}H_{96}N_{12}O_{12}$Na) 1199.7168, found 1199.7227.

General procedure for the synthesis of the Ac-Glu-Tyr-OH dipeptides: Wang resin (2.5 g, 100-200 mesh particle size, extent of labelling: 1.1 mmol/g loading, 2.75 mmol) was suspended in 1 : 1 DCM : DMF mixture (40 ml). In a separate flask, Fmoc-Tyr(*t*Bu)-OH (11 mmol) was dissolved in 15 ml of dry DMF, to which solution hydroxybenzotriazole (HOBt, 1.68 g, 11 mmol) was added. This solution was added to the resin suspension, followed by DMAP (50 mg, 0.40 mmol). After 10 minutes *N,N'*-Diisopropylcarbodiimide (DIC, 1.70 ml, 11 mmol) was added, and the reaction mixture shaken during 4 hours. The resin was filtered, washed with DCM and DMF (*ca.* 150 ml total volume) and dried. It was suspended in 1 : 4 piperidine : DMF and shaken during 10 minutes. Resin was filtered and the treatment repeated. It was filtered again and washed with DMF, *i*PrOH and DCM. Presence of primary amine was assessed by reacting a small amount of resin with 2,4,6-trinitrobenzensulfonic acid (TNBS) in presence of DIPEA (red color). Resin was suspended in 1 : 1 DCM : DMF mixture (50 ml). In a separate flask, Fmoc-Glu(*t*Bu)-OH (2.05 g, 6.87 mmol), HATU (2.61 g, 6.87 mmol) and DIPEA (2.40 ml, 13.7 mmol) were dissolved in 20 ml of dry DMF, and this solution was added to the resin suspension. The reaction mixture was shaken during 16 hours. It was filtered and washed with DMF and DCM (*ca.* 150 ml total volume) and dried. It was suspended in 1 : 4 piperidine : DMF and shaken during 10 minutes. Resin was filtered and the treatment repeated. It was filtered again and washed with DMF, *i*PrOH and DCM. Presence of primary amine was assessed by reacting a small amount of resin with 2,4,6-trinitrobenzensulfonic acid (TNBS) in presence of DIPEA (red color). Resin was suspended in DCM (40 ml); acetyl anhydride (1.30 ml, 13.7 mmol) and DIPEA (2.60 ml, 15 mmol) were added to the resin suspension, and the reaction mixture shaken during 4 hours. The resin was filtered, washed with plenty of DCM and dried. It was suspended in 50 : 50 : 1 TFA : DCM : TES mixture (80 ml). It was stirred at room temperature during 2 hours. The resin was filtered and washed with plenty of DCM. The filtrate was concentrated. The residue was washed several times with hexane and diethyl ether until a pale yellow solid was obtained. The crude Ac-EY-OH dipeptide was purified by reversed-phase column chromatography (mobile phase: 1% to 20% MeCN in water). Pure Ac-EY-OH dipeptides were obtained as white solids, overall typical yield was *ca.* 35%.

NMR titration procedures

The titrations were performed with the cage receptors as free amines. Stock solutions of the cages were prepared by weighting the corresponding amount of the receptor and reaching a final concentration around 1 mM in in the chosen solvents mixture (a) 2 : 1 CD_3CN : H_2O; b) 1 : 1 CD_3CN : H_2O; c) 6.4 : 1 CD_3CN : H_2O).

Stock solutions of the titrant containing 20-45 mM dipeptide were prepared by dissolving the dipeptides in the stock solution of the corresponding cage, thus maintaining the concentration of the cage constant during the titration experiment. The stock solution of the cage was introduced in a NMR tube and the ^{1}H NMR spectrum (500 MHz, 298 K) was acquired using the water excitation sculpting (DPFGSE) sequence[27] from the CHEMPACK library of Agilent VnmrJ32 software. Then volumes of the stock solution of the titrant were added and the ^{1}H NMR spectrum recorded after each addition. Different signals shifted upon addition of the dipeptides, and their shifts were fitted to a 1 : 1 receptor : substrate model using HypNmr 2008 version 4.0.71 software.[28]

ESI-MS experiments

A Q-TOF Premier (Waters) mass spectrometer with an electrospray source operating in the V-mode was used. The drying gas as well as the cone gas was nitrogen at a flow of 300 Lh^{-1} and 30 Lh^{-1}, respectively. The temperature of the source block was set to 100 °C and the desolvation temperature was set to 150 °C. A capillary voltage of 3.5 kV was used in the positive scan mode and the cone voltage was adjusted typically to *U*c = 10 V to control the extent of fragmentation. Mass calibration was performed by using NaI solutions in isopropanol:water (1:1) from m/z 50–3000.

For the application of the enantiomer-labelled method, the isotopically labelled d^3-Ac-L-Glu-L-Tyr-OH guest was thought as an inexpensive, easy to prepare pseudostereoisomer to distinguish the diastereomeric host-guest complex ions in the corresponding ESI mass spectra. The labelled dipeptide was also synthesized by conventional solid-phase peptide synthesis from *N*-Fmoc-amino acids, but capping the *N*-terminus with

CD_3COCl in basic medium. Despite the analytical data of d^3-Ac-L-Glu-L-Tyr-OH indicate that epimerisation did not occur during the synthesis, the ESI-MS analysis revealed that samples indeed comprise a mixture of d^3-Ac-L-Glu-L-Tyr-OH (~90 %) and d^2-Ac-L-Glu-L-Tyr-OH (~10 %). The NMR analysis of the sample indicated that the 10% H/D exchange occurred in the final acetylation step (most likely through the partial formation of the corresponding deuterated ketene and its subsequent reaction with the amino group during the capping) thus locating the residual H/D exchanged protons in the methyl of the *N*-terminal acetyl group. However, this does not hamper establishing the chiral recognition trends based on ESI-MS. Because acetyl group is far from the major binding sites of the host, it is expected that these labels have little or no effect on the binding selectivity measured by ESI-MS. In addition, it is also unlikely that this minor modification affects the ionization efficiencies of the non-labelled and labelled complexes. Control experiments were performed using equimolecular mixtures of d^3-Ac-L-Glu-L-Tyr-OH and non-labelled Ac-L-Glu-L-Tyr-OH and overall revealed identical binding affinities (see table 1 and figure S47). Rare examples of pronounced effects on the binding abilities estimated by ESI-MS upon deuteration have been reported for cinchonane-type chiral hosts and model chiral acids as guests.[29]

Sample solutions were prepared as follows: to equimolar 1 x 10^{-3} M acetonitrile solutions of d^3-Ac-L-Glu-L-Tyr-OH and one of the different stereoisomers of the Ac-Glu-Tyr dipeptide was added one equivalent of **CySer** (or **CyThr**) host. The resulting mixture was diluted with CH_3CN:H_2O (2:1) to a final 5 x 10^{-5} M concentration and analyzed by positive ESI-MS. Since one host–guest complex is labelled with only a CD_3 group, the two diastereomeric host–guest complexes are partially overlapped as illustrated in Fig. 5. Nevertheless, chiral recognition can be readily confirmed from the relative peak intensities of both complexes. For collision induced dissociation (CID) experiments, the complete isotopic envelope of the species of interest was mass-selected with the first quadrupole (isolation width *ca.* 0.5 Da), interacted with argon in the T-wave collision cell while analyzing the ionic fragments with the TOF analyzer. The collision energy ($CE_{laboratory}$) was systematically stepped in the E_{lab} = 1-10 eV range. CID mass spectra are shown at $CE_{laboratory}$ = 5 eV to clearly illustrate the distinctive product ion abundances of the investigated ions.

Molecular modelling

All the theoretical calculations were performed using Spartan '06 program working on a Dell workstation and with the MMFF force field. The complex between **CySer** cage and the Ac-L-Glu-L-Tyr-OH dipeptide was manually built considering the previously found experimental results.[14] The geometry thus obtained was subjected to Monte Carlo searches with MMFF minimizations (in the gas phase), starting from different conformations of the Ser side chains of the host. The distances between the aromatic rings of the host and the Tyr residue of the guest were constrained to less than 4.5 Å in order to reflect the NOE contact observed in the corresponding ROESY experiment. The lowest energy minimum thus obtained is shown in Fig. 1. For the complexes formed with the other isomers of the dipeptides, we manually inverted the respective configurations of the Cα chiral centres from the LL-complex and repeated the process rendering the minima shown in Fig. 7. The same procedure was used for the **CyThr** receptor (See results in the ESI). The alignment of the obtained minima was carried out also with the same software. The final representation of the structures was prepared with ViewerLite 4.2 for aesthetical reasons.

Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO; CTQ2012-38543-C03 project) and Generalitat de Catalunya (AGAUR, 2014 SGR 231). E.F. thanks CSIC and the European Social Fund for personal financial support (JAE-doc program). We thank SCIC of the Universitat Jaume I for providing us with mass spectrometric facilities and Dr. Yolanda Pérez (IQAC) for her helpful assistance with the acquisition of advanced NMR experiments.

Notes and references

1. a) W. C. Still, *Acc. Chem. Res.*, 1996, **29**, 155; b) P. D. Henley and J. D. Kilburn, *Chem. Commun.*, 1999, 1335; c) J. Dowden, P. D. Edwards, S. S. Flack and J. D. Kilburn, *Chem. Eur. J.*, 1999, **5**, 79; d) M. W. Peczuh and A. D. Hamilton, *Chem. Rev.*, 2000, **100**, 2479; e) S. Tashiro, M. Tominaga, M. Kawano, B. Therrien, T. Ozeki and M. Fujita, *J. Am. Chem. Soc.*, 2005, **127**, 4546; f) M. Wehner, D. Janssen, G. Schäfer and T. Schrader, *Eur. J. Org. Chem.*, 2006, 138; g) T. Gersthagen, C. Schmuck and T. Schrader, *Supramol. Chem.*, 2010, **22**, 853; h) A. Späth and B. König, *Tetrahedron*, 2010, **66**, 6019; i) S. Zhang, L. Han, C.-g. Li, J. Wang, W. Wang, Z. Yuan and X. Gao, *Tetrahedron*, 2012, **68**, 2357; j) D. Maity and C. Schmuck, in *Monographs in Supramolecular Chemistry*, RSC Publishing, 2015, pp. 326-368.
2. a) C. Schmuck and L. Geiger, *J. Am. Chem. Soc.*, 2004, **126**, 8898; b) C. Schmuck, *Coord. Chem. Rev.*, 2006, **250**, 3053; c) C. Schmuck and L. Hernandez-Folgado, *Org. Biomol. Chem.*, 2007, **5**, 2390; d) F. Biedermann, U. Rauwald, M. Cziferszky, K. A. Williams, L. D. Gann, B. Y. Guo, A. R. Urbach, C. W. Bielawski and O. A. Scherman, *Chem. Eur. J.*, 2010, **16**, 13716; e) T. Gersthagen, J. Hofmann, F.-G. Klärner, C. Schmuck and T. Schrader, *Eur. J. Org. Chem.*, 2013, 1080; f) F. Biedermann and W. M. Nau, *Angew. Chem. Int. Ed.*, 2014, **53**, 5694; g) S. Yapar, M. Oikonomou, A. H. Velders and S. Kubik, *Chem. Commun.*, 2015, **51**, 14247.
3. a) L. M. Heitmann, A. B. Taylor, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2006, **128**, 12574; b) M. V. Rekharsky, H. Yamamura, Y. H. Ko, N. Selvapalam, K. Kim and Y. Inoue, *Chem. Commun.*, 2008, 2236; c) J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2011, **133**, 8810; d) L. A. Logsdon, C. L. Schardon, V. Ramalingam, S. K. Kwee and A. R. Urbach, *J. Am. Chem. Soc.*, 2011, **133**, 17087; e) O. A. Ali, E. M. Olson and A. R. Urbach, *Supramol. Chem.*, 2013, **25**, 863; f) L. A. Logsdon and A. R. Urbach, *J. Am. Chem. Soc.*, 2013, **135**, 11414; g) L. C. Smith, D. G. Leach, B. E. Blaylock, O. A. Ali and A. R. Urbach, *J. Am. Chem. Soc.*, 2015, **137**, 3663.

4. a) M. Kubo, E. Nashimoto, T. Tokiyo, Y. Morisaki, M. Kodama and H. Hioki, *Tetrahedron Lett.*, 2006, **47**, 1927; 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) F. Sansone, L. Baldini, A. Casnati and R. Ungaro, *New J. Chem.*, 2010, **34**, 2715; d) A. D. Stancu, H.-J. Buschmann and L. Mutihac, *J. Incl. Phenom. Macrocycl. Chem.*, 2012, **75**, 1.
5. a) H. Yamamura, M. V. Rekharsky, Y. Ishihara, M. Kawai and Y. Inoue, *J. Am. Chem. Soc.*, 2004, **126**, 14224; b) Y. Liu, Y.-W. Yang, Y. Chen and F. Ding, *Biorg. Med. Chem.*, 2005, **13**, 963; c) Y. Zhao, J. Gu, Y. C. Yang, H. Y. Zhu, R. Huang and B. Jing, *J. Mol. Struct.*, 2009, **930**, 72; d) H. S. Christensen, B. W. Sigurskjold, T. G. Frihed, L. G. Marinescu, C. M. Pedersen and M. Bols, *Eur. J. Org. Chem.*, 2011, 5279; e) M. Altarsha, V. Yeguas, F. Ingrosso, R. Lopez and M. F. Ruiz-Lopez, *J. Phys. Chem. B*, 2013, **117**, 3091.
6. a) M. F. Cristofaro and A. R. Chamberlin, *J. Am. Chem. Soc.*, 1994, **116**, 5089; b) M. V. Rekharsky, H. Yamamura, M. Kawai and Y. Inoue, *J. Org. Chem.*, 2003, **68**, 5228; c) C. Bombelli, S. Borocci, F. Lupi, G. Mancini, L. Mannina, A. L. Segre and S. Viel, *J. Am. Chem. Soc.*, 2004, **126**, 13354; d) H. Imai, H. Munakata, Y. Uemori and N. Sakura, *Inorg. Chem.*, 2004, **43**, 1211; e) K.-H. Chang, J.-H. Liao, C.-T. Chen, B. K. Mehta, P.-T. Chou and J.-M. Fang, *J. Org. Chem.*, 2005, **70**, 2026; f) O. Cruciani, S. Borocci, R. Lamanna, G. Mancini and A. L. Segre, *Tetrahedron: Asymmetry*, 2006, **17**, 2731; g) C. Schmuck and P. Wich, *Angew. Chem. Int. Ed.*, 2006, **45**, 4277; h) A. M. Castilla, M. Morgan Conn and P. Ballester, *Beil. J. Org. Chem.*, 2010, **6**, 5; i) J. Ren, J. Wang, J. Wang and E. Wang, *Chem. Eur. J.*, 2013, **19**, 479.
7. a) J. I. Hong, S. K. Namgoong, A. Bernardi and W. C. Still, *J. Am. Chem. Soc.*, 1991, **113**, 5111; b) M. Conza and H. Wennemers, *J. Org. Chem.*, 2002, **67**, 2696; c) S. Kubik, in *Artificial Receptors for Chemical Sensors*, Wiley-VCH, 2010, , pp. 135-167; d) S. Niebling, H. Y. Kuchelmeister, C. Schmuck and S. Schlücker, *Chem. Sci.*, 2012, **3**, 3371.
8. a) M. Nieto and H. R. Perkins, *Biochem. J*, 1971, **123**, 789; b) M. Nieto and H. R. Perkins, *Biochem. J*, 1971, **123**, 773; c) H. Molinari, A. Pastore, L. Y. Lian, G. E. Hawkes and K. Sales, *Biochemistry*, 1990, **29**, 2271.
9. S. V. Luis and I. Alfonso, *Acc. Chem. Res.*, 2014, **47**, 112.
10. a) I. Martí, J. Rubio, M. Bolte, M. I. Burguete, C. Vicent, R. Quesada, I. Alfonso and S. V. Luis, *Chem. Eur. J.*, 2012, **18**, 16728; b) I. Martí, M. Bolte, M. I. Burguete, C. Vicent, I. Alfonso and S. V. Luis, *Chem. Eur. J.*, 2014, **20**, 7458.
11. I. Alfonso, M. Bolte, M. Bru, M. I. Burguete, S. V. Luis and C. Vicent, *Org. Biomol. Chem.*, 2010, **8**, 1329.
12. a) A. Moure, S. V. Luis and I. Alfonso, *Chem. Eur. J.*, 2012, **18**, 5496. b) V. Martí-Centelles, M. D. Pandey, M. I. Burguete and S. V. Luis, *Chem. Rev.*, 2015, **115**, 8736.
13. a) S. Kubik, *Chem. Soc. Rev.*, 2009, **38**, 585; b) R. B. Elmes and K. A. Jolliffe, *Chem. Commun.*, 2015, **51**, 4951.
14. E. Faggi, A. Moure, M. Bolte, C. Vicent, S. V. Luis and I. Alfonso, *J. Org. Chem.*, 2014, **79**, 4590.
15. a) Y. Kuroda, Y. Kato, T. Higashioji, J.-y. Hasegawa, S. Kawanami, M. Takahashi, N. Shiraishi, K. Tanabe and H. Ogoshi, *J. Am. Chem. Soc.*, 1995, **117**, 10950; b) X. Mei and C. Wolf, *J. Am. Chem. Soc.*, 2004, **126**, 14736; c) J. Zhao, T. M. Fyles and T. D. James, *Angew. Chem. Int. Ed.*, 2004, **43**, 3461; d) S. Shinoda, T. Okazaki, T. N. Player, H. Misaki, K. Hori and H. Tsukube, *J. Org. Chem.*, 2005, **70**, 1835; e) M. V. Rekharsky, H. Yamamura, C. Inoue, M. Kawai, I. Osaka, R. Arakawa, K. Shiba, A. Sato, Y. H. Ko, N. Selvapalam, K. Kim and Y. Inoue, *J. Am. Chem. Soc.*, 2006, **128**, 14871; f) X. Li, M. Tanasova, C. Vasileiou and B. Borhan, *J. Am. Chem. Soc.*, 2008, **130**, 1885; g) B. Altava, D. S. Barbosa, M. I. Burguete, J. Escorihuela and S. V. Luis, *Tetrahedron: Asymmetry*, 2009, **20**, 999; h) A. Sirikulkajorn, T. Tuntulani, V. Ruangpornvisuti, B. Tomapatanaget and A. P. Davis, *Tetrahedron*, 2010, **66**, 7423; i) T. P. Quinn, P. D. Atwood, J. M. Tanski, T. F. Moore and J. F. Folmer-Andersen, *J. Org. Chem.*, 2011, **76**, 10020; j) S. Yu, W. Plunkett, M. Kim and L. Pu, *J. Am. Chem. Soc.*, 2012, **134**, 20282; k) F. Ulatowski and J. Jurczak, *J. Org. Chem.*, 2015, **80**, 4235.
16. H. J. Schneider, P. Agrawal and A. K. Yatsimirsky, *Chem. Soc. Rev.*, 2013, **42**, 6777.
17. a) M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren and R. A. Friesner, *J. Am. Chem. Soc.*, 1997, **119**, 5908; b) E. F. Strittmatter and E. R. Williams, *J. Phys. Chem. A*, 2000, **104**, 6069; c) C. F. Rodriquez, G. Orlova, Y. Guo, X. Li, C.-K. Siu, A. C. Hopkinson and K. W. M. Siu, *J. Phys. Chem. B*, 2006, **110**, 7528; d) W. Brandt, T. Herberg and L. Wessjohann, *Biopolymers*, 2011, **96**, 651.
18. a) D. W. Boykin, *NMR Spectroscopy in Organic Chemistry*, CRC press, 1990; b) T. D. Claridge, *High-resolution NMR techniques in organic chemistry*, Newnes, 2008.
19. a) M. Pons, *NMR in supramolecular chemistry*, Springer Science & Business Media, 1999; b) M. Pons and O. Millet, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2001, **38**, 267; c) A. Pastor and E. Martínez-Viviente, *Coord. Chem. Rev.*, 2008, **252**, 2314.
20. a) K. Hirose, *J. Incl. Phenom.*, 2001, **39**, 193; b) P. Thordarson, *Chem. Soc. Rev.*, 2011, **40**, 1305; c) J. García, L. G. Martins and M. Pons, in *Supramolecular Chemistry*, John Wiley & Sons, Ltd, 2012.
21. a) S. Braun, W. E. Raymond and E. Racker, *J. Biol. Chem.*, 1984, **259**, 2051; b) P. Cohen, *Nat. Rev. Drug Discov.*, 2002, **1**, 309; c) J. F. Glickman and T. Chen, *Assay development for protein kinases and phosphatases*, CRC Press: Boca Raton, FL, 2010.
22. M. Rekharsky and Y. Inoue, in *Supramolecular Chemistry*, John Wiley & Sons, Ltd, 2012.
23. E. Klein, Y. Ferrand, N. P. Barwell and A. P. Davis, *Angew. Chem. Int. Ed.*, 2008, **47**, 2693.
24. a) A. Di Tullio, S. Reale and F. De Angelis, *J. Mass Spectrom.*, 2005, **40**, 845; b) B. Baytekin, H. T. Baytekin and C. A. Schalley, *Org. Biomol. Chem.*, 2006, **4**, 2825; c) Z. Chen and S. G. Weber, *TRAC-Trend. Anal. Chem.*, 2008, **27**, 738; d) C. Bich, S. Baer, M. C. Jecklin and R. Zenobi, *J. Am. Soc. Mass. Spectrom.*, 2010, **21**, 286.
25. a) M. Sawada, *Mass Spectrom. Rev.*, 1997, **16**, 73; b) K. A. Schug and W. Lindner, *J. Pep. Sci.*, 2005, **28**, 1932; c) K. A. Schug, *Comb. Chem. High T. Scr.*, 2007, **10**, 301; d) L. Wu and F. G. Vogt, *J. Pharm. Biomed. Anal.*, 2012, **69**, 133; e) H. Awad and A. El-Aneed, *Mass Spectrom. Rev.*, 2013, **32**, 466.
26. M. Sawada, Y. Takai, H. Yamada, J. Nishida, T. Kaneda, R. Arakawa, M. Okamoto, K. Hirose, T. Tanaka and K. Naemura, *J. Chem. Soc., Perkin Trans. 2*, 1998, 701.
27. T. L. Hwang and A. J. Shaka, *J. Magn. Reson., Ser A*, 1995, **112**, 275.
28. a) C. Frassineti, S. Ghelli, P. Gans, A. Sabatini, M. S. Moruzzi and A. Vacca, *Anal. Biochem.*, 1995, **231**, 374; b) C. Frassineti, L. Alderighi, P. Gans, A. Sabatini, A. Vacca and S. Ghelli, *Anal. Bioanal. Chem.*, 2003, **376**, 1041.
29. K. A. Schug, N. M. Maier and W. Lindner, *J. Mass Spectrom.*, 2006, **41**, 157.