Electronic Supplementary Material (ESI) for Analyst.


# Supplementary Information
# For
# Pseudopeptidic Host Adaptation in Peptide Recognition Unveiled by Ion Mobility Mass Spectrometry

Lucia Tapia,[a] Yolanda Pérez,[b] Jordi Solà,[a] Santiago V. Luis,[c] Ignacio Alfonso, [a,*] Cristian Vicent[d,*]

a. Department of Biological Chemistry, Institute for Advanced Chemistry of Catalonia, IQAC-CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain. E-mail: ignacio.alfonso@iqac.csic.es

b. NMR Facility, Institute for Advanced Chemistry of Catalonia, IQAC-CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain.

c. Departamento de Química Inorgánica y Orgánica, Universitat Jaume I, Av. de Vicent Sos Baynat s/n, 12071 Castellón, Spain

d. Servei Central d'Instrumentació Científica (SCIC), Universitat Jaume I, Avda. Sos Baynat s/n, 12006 Castellón, Spain; E-mail: barrera@uji.es

## Table of contents:

**Potentiometric titrations of CyOrn and CyLys**

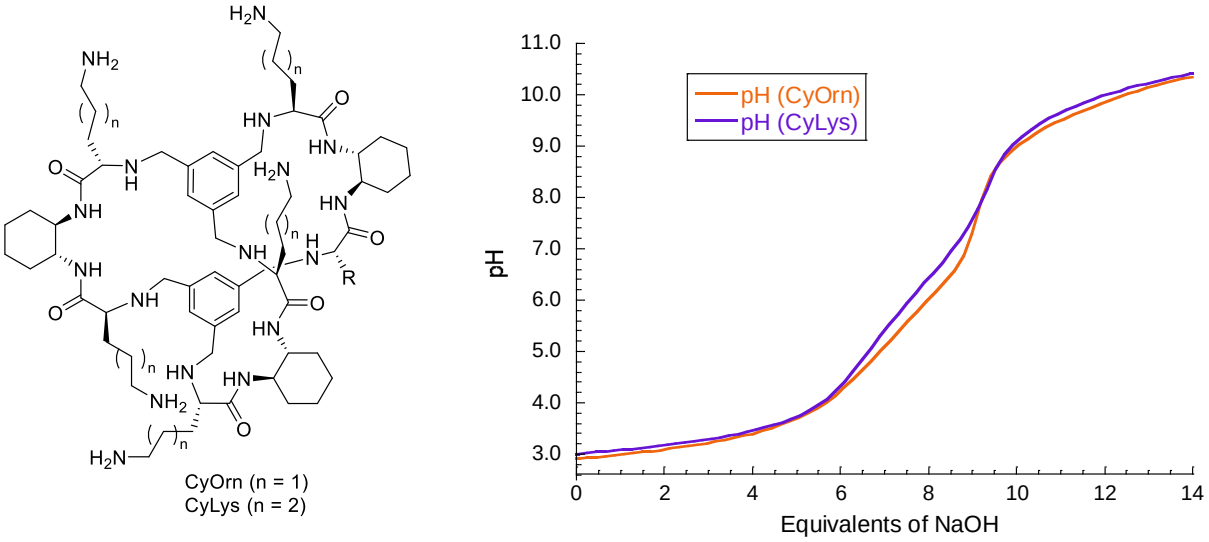


**Figure S1**. Superposition of the potentiometric titration curves for the HCl salts (with a 20% excess of HCl) of **CyLys** and **CyOrn** cages (~0.5 mM cage in 150 mM NaCl titrated with 0.1 M NaOH). These systems displayed too many protonation / deprotonation equilibria for a reliable fitting of pKa values. The CyLys curve is above CyOrn curve meaning that CyLys is on average more basic (more charged at a given pH). Starting from fully protonated species ($12H^+$ with 20% excess HCl, meaning 14 equivalents of $H^+$), 8-10 equiv. of NaOH are needed to reach pH 6-8, thus the number of protonation sites around neutral pH would be 4-6 $H^+$.

## NMR studies of CyOrn at different pH values

NMR experiments were acquired at 500 MHz using a Bruker Avance III spectrometer with a $^{1}H/^{15}N/^{13}C$ TCl cryoprobe and *z* field gradient coil. NMR experiments of CyOrn cage at different pHs were conducted at 25°C in 50 mM Tris-d11/100% $D_2O$. All data were processed with MNova software package (Mestrelab Research S.L., Santiago de Compostela, Spain). Proton NMR experiments were acquired using standard pulse sequence zggpw5 (proton with water suppression using Watergate W5 with gradients) from Bruker Topspin 3.6 (Bruker Biospin GmbH, Ettlingen, Germany).

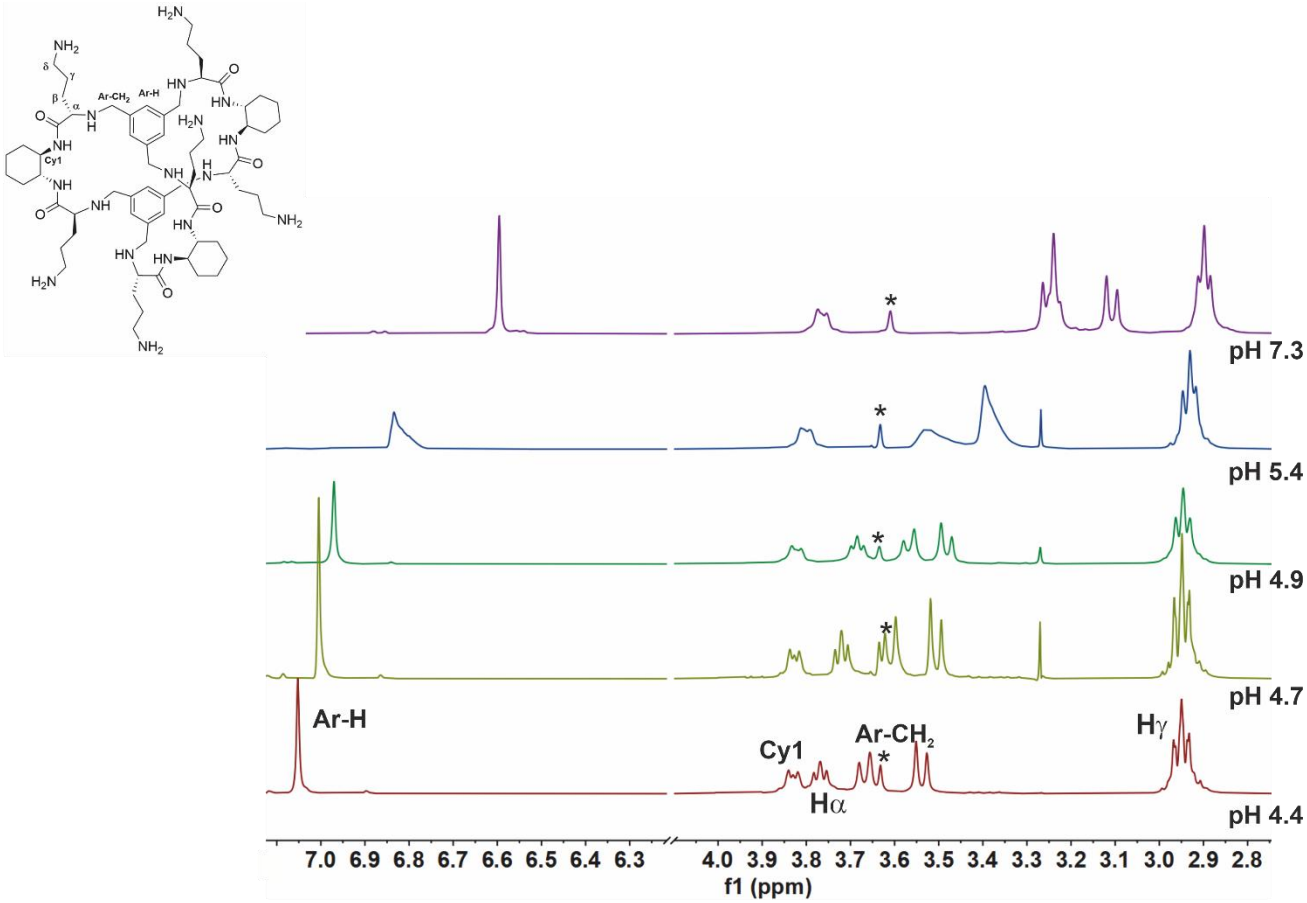


**Figure S2.** $^{1}H$ 1D NMR spectra of **CyOrn** cage at different pH values, indicated at each trace. The signal with an asterisk (*) corresponds to a residual proton resonance from Tris-d11 buffer. The spectra from acidic to neutral pH show up-field chemical shift variation of the signals within the cage core (Ar-H, Ar-$CH_2$ and Hα) implying a progressive deprotonation of the secondary ammonium groups. This observation confirms that, at neutral pH, the positive charges are mainly located at the primary amine nitrogen atoms of the side chains, as expected from the corresponding individual pKa values and also to minimize electrostatic repulsions.

**Mass Spectrometry characterization of the protonated species**

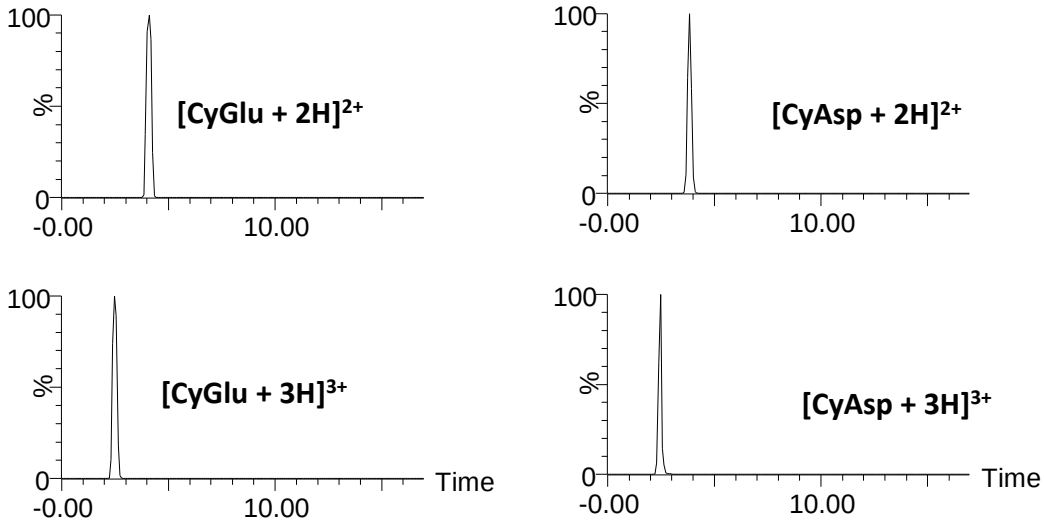


**Figure S3**. Arrival time distribution for the species generated from aqueous solution of **CyGlu** (left) and **CyAsp** (right).

## Fluorescence titrations for the determination of the EYE-cages binding constants

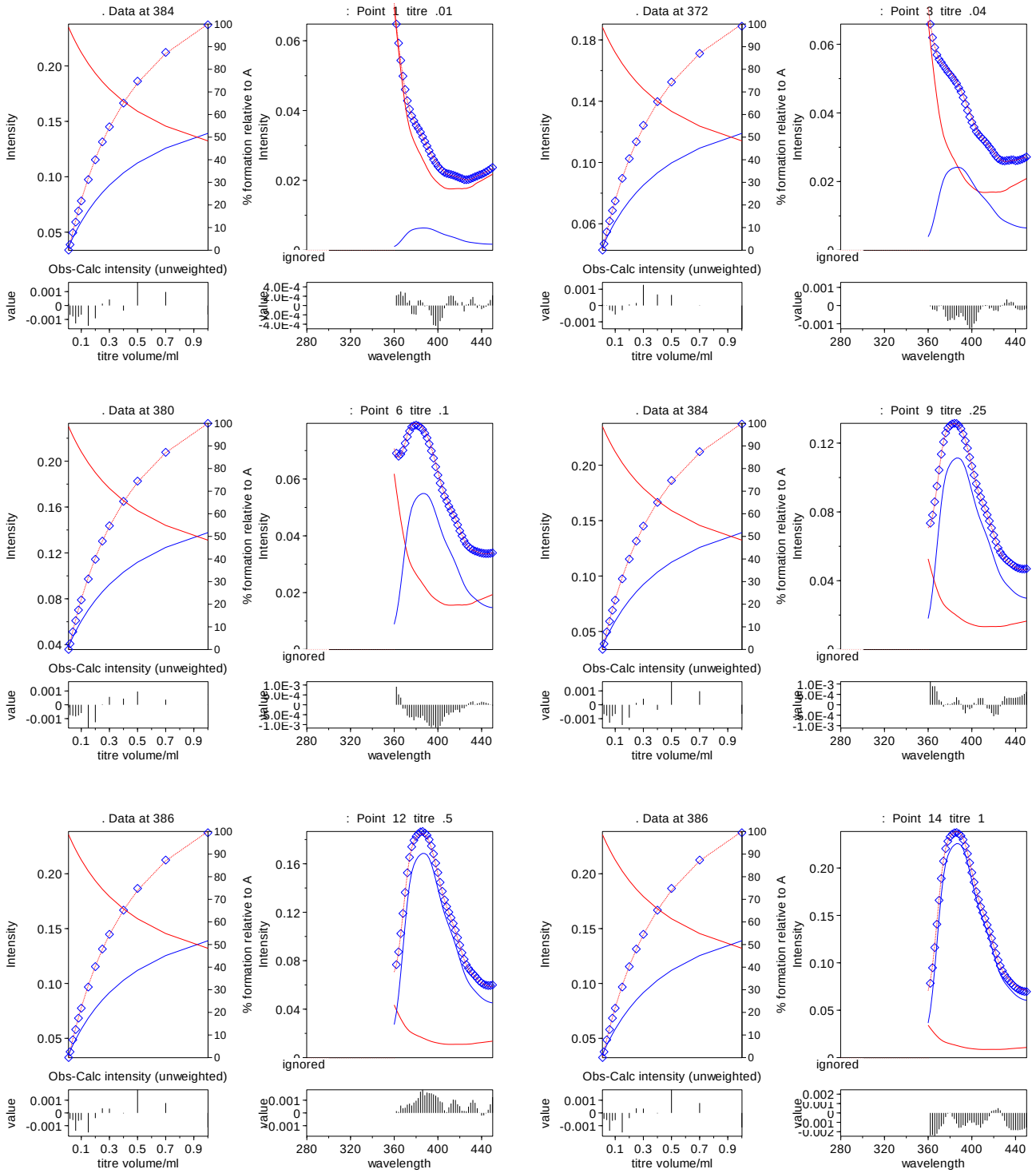


**Figure S4.** Fitting of representative points for the titration of EYE with **CyOrn** (20 ˚C, 50 mM Tris, pH 7.3, [EYE] = $2 \cdot 10^{-4}$ M, [CyOrn] = 0-$1.5 \cdot 10^{-3}$ M). On the left side of each panel we show the titration curve (experimental points as blue symbols) with the fitting (red dashed line) and the species distribution (free EYE and the 1:1 complex in red and blue lines, respectively); while on the right side of each panel we include the fluorescence emission spectra (experimental as blue symbols, best fit as red dashed lines) as well as the emission spectra for the two species (free EYE and the 1:1 complex in red and blue lines, respectively) obtained from the fitting. The corresponding residuals are plotted at the bottom of each graphic.

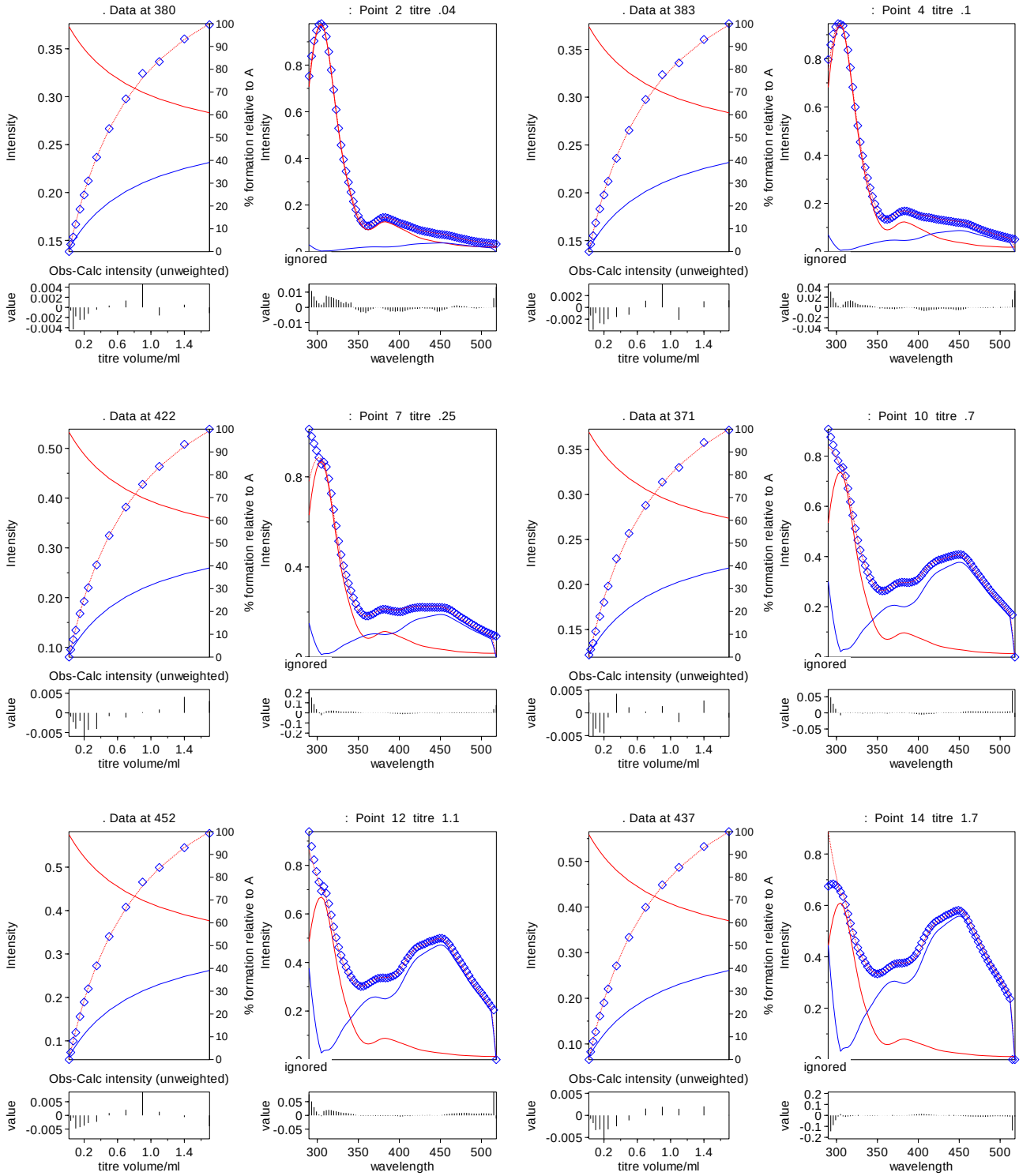


**Figure S5.** Fitting of representative points for the titration of EYE with **CyLys** (20 ˚C, 50 mM Tris, pH 7.3, [EYE] = $4 \cdot 10^{-5}$ M, [CyLys] = 0-$1 \cdot 10^{-3}$ M). On the left side of each panel we show the titration curve (experimental points as blue symbols) with the fitting (red dashed line) and the species distribution (free EYE and the 1:1 complex in red and blue lines, respectively); while on the right side of each panel we include the fluorescence emission spectra (experimental as blue symbols, best fit as red dashed lines) as well as the emission spectra for the two species (free EYE and the 1:1 complex in red and blue lines, respectively) obtained from the fitting. The corresponding residuals are plotted at the bottom of each graphic.

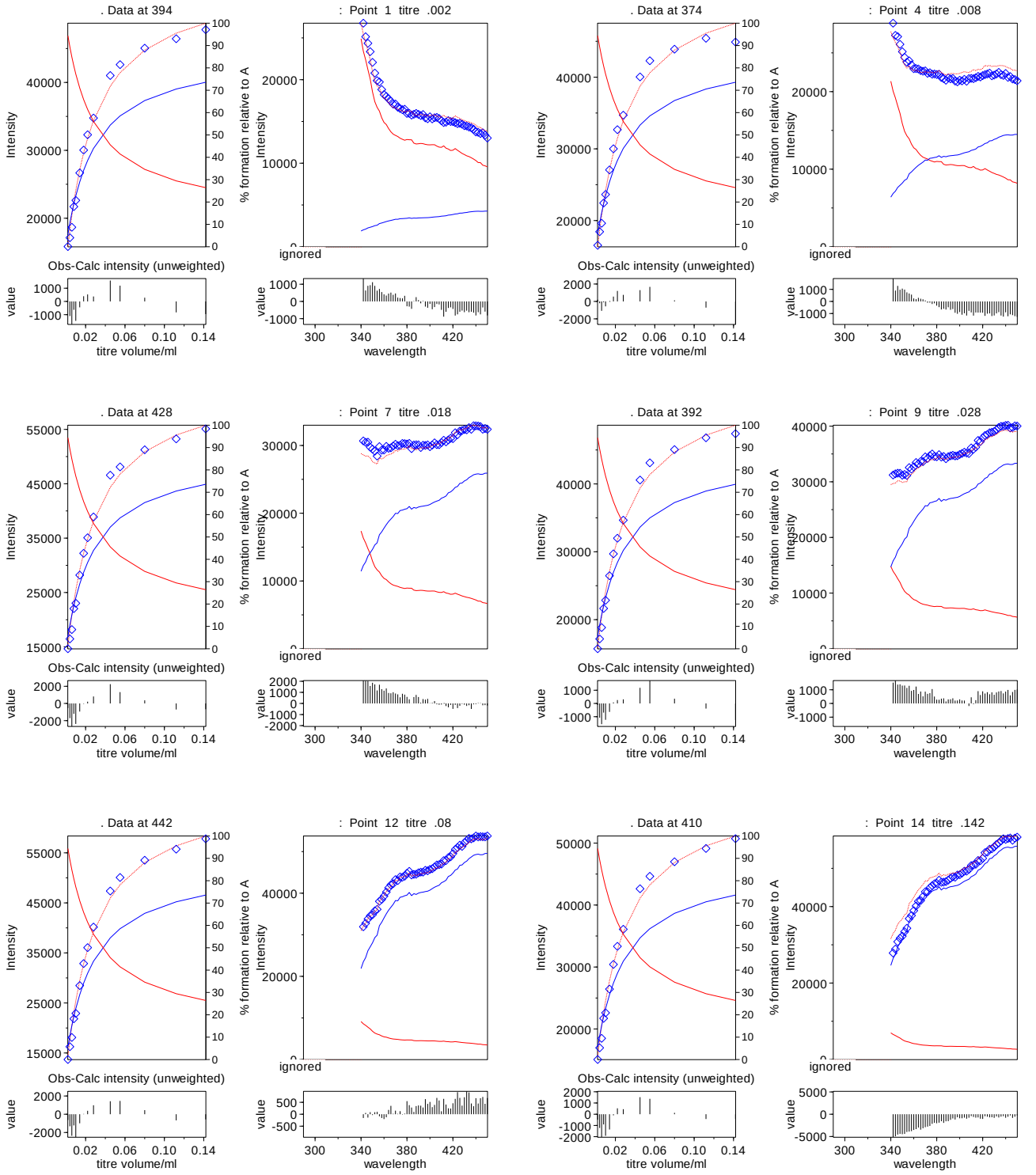


**Figure S6.** Fitting of representative points for the titration of EYE with **CyHis** (20 ˚C, 50 mM Tris, pH 7.3, [EYE] = $2 \cdot 10^{-5}$ M, [CyHis] = 0-$1 \cdot 10^{-3}$ M). On the left side of each panel we show the titration curve (experimental points as blue symbols) with the fitting (red dashed line) and the species distribution (free EYE and the 1:1 complex in red and blue lines, respectively); while on the right side of each panel we include the fluorescence emission spectra (experimental as blue symbols, best fit as red dashed lines) as well as the emission spectra for the two species (free EYE and the 1:1 complex in red and blue lines, respectively) obtained from the fitting. The corresponding residuals are plotted at the bottom of each graphic.

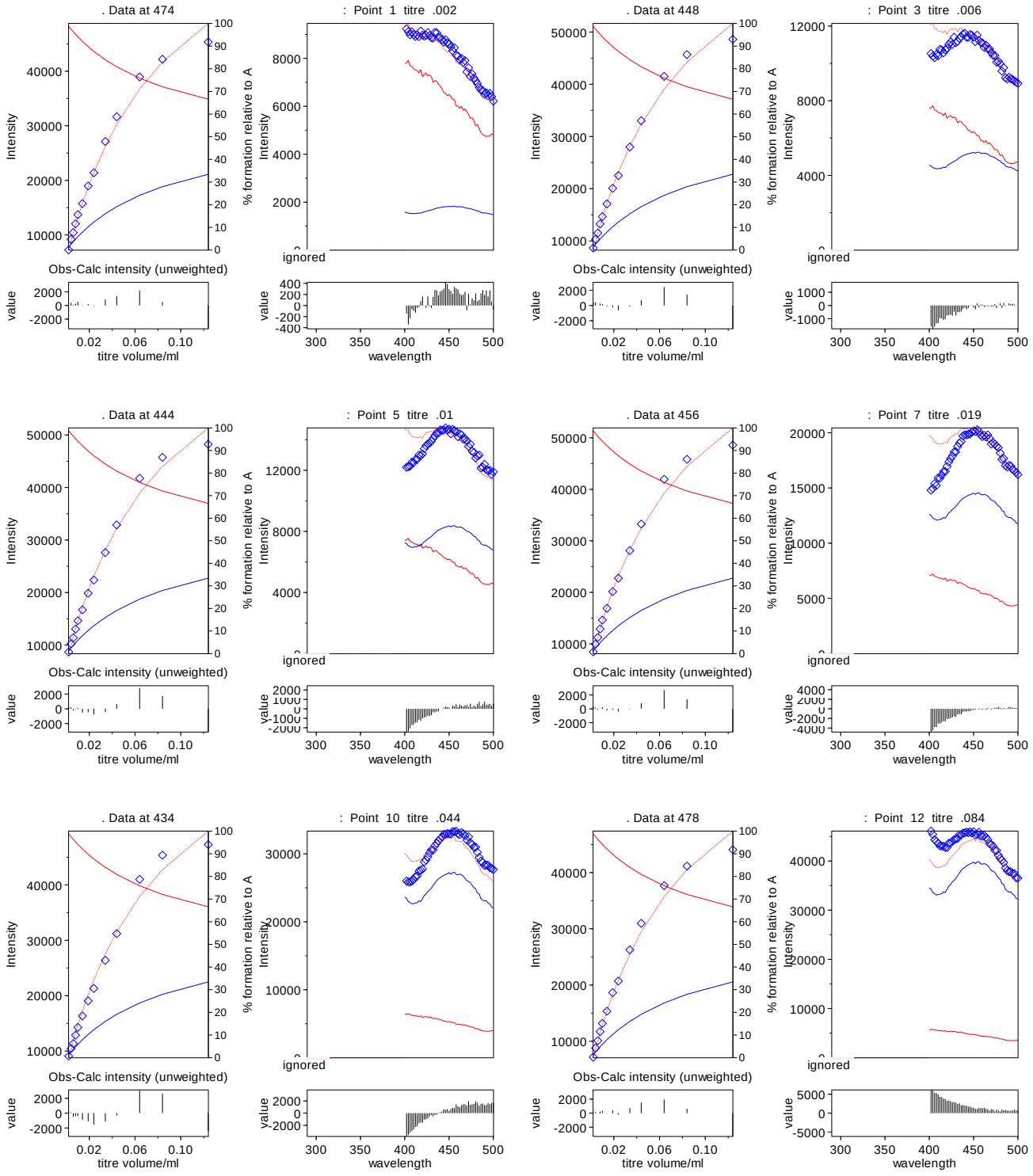


**Figure S7.** Fitting of representative points for the titration of EYE with **CyAsp** (20 ˚C, 50 mM Tris, pH 7.3, [EYE] = $1 \cdot 10^{-5}$ M, [CyAsp] = 0-$1 \cdot 10^{-3}$ M). On the left side of each panel we show the titration curve (experimental points as blue symbols) with the fitting (red dashed line) and the species distribution (free EYE and the 1:1 complex in red and blue lines, respectively); while on the right side of each panel we include the fluorescence emission spectra (experimental as blue symbols, best fit as red dashed lines) as well as the emission spectra for the two species (free EYE and the 1:1 complex in red and blue lines, respectively) obtained from the fitting. The corresponding residuals are plotted at the bottom of each graphic.

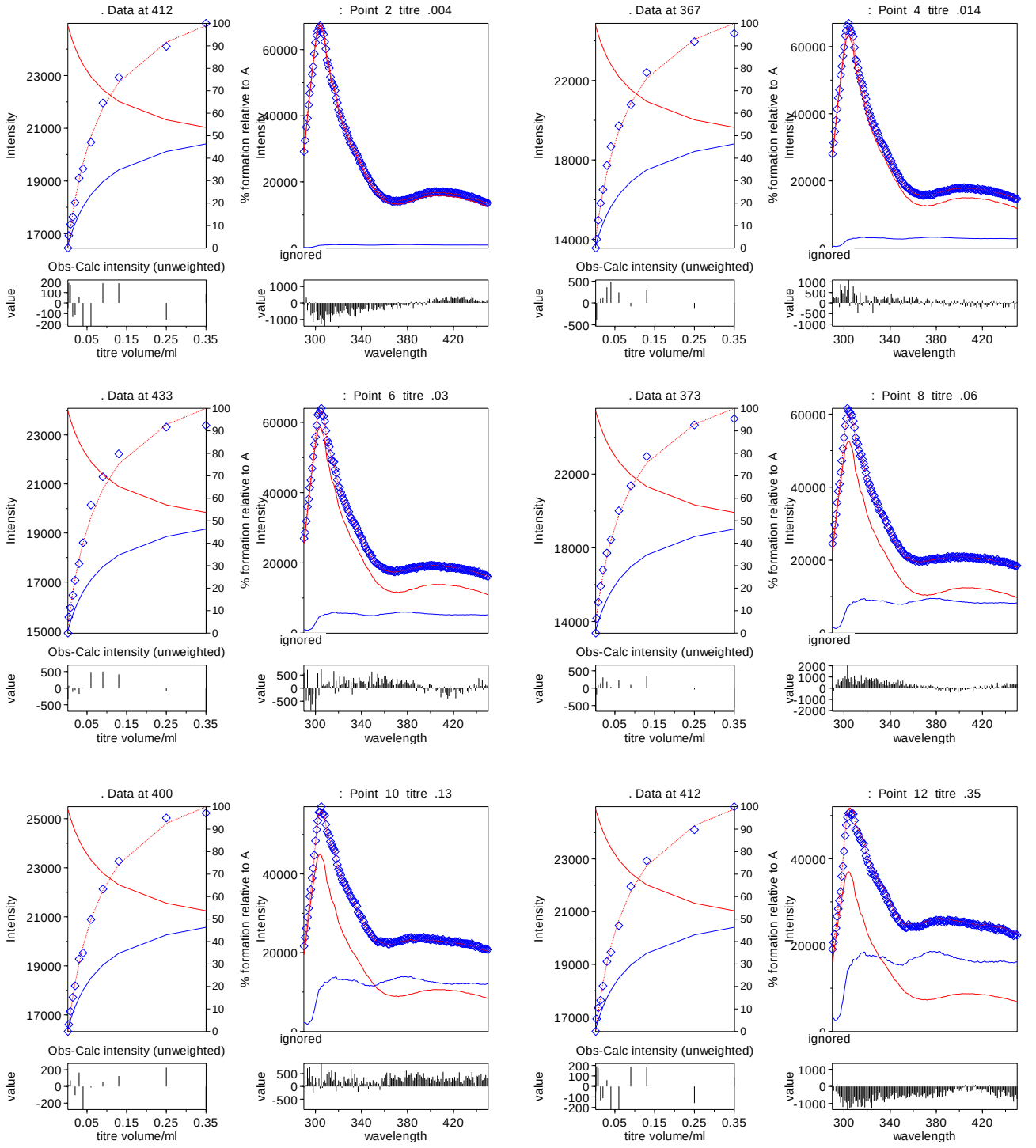


**Figure S8.** Fitting of representative points for the titration of EYE with **CyGlu** (20 ˚C, 50 mM Tris, pH 7.3, [EYE] = $2 \cdot 10^{-5}$ M, [CyGlu] = 0-1 $\cdot 10^{-3}$ M). On the left side of each panel we show the titration curve (experimental points as blue symbols) with the fitting (red dashed line) and the species distribution (free EYE and the 1:1 complex in red and blue lines, respectively); while on the right side of each panel we include the fluorescence emission spectra (experimental as blue symbols, best fit as red dashed lines) as well as the emission spectra for the two species (free EYE and the 1:1 complex in red and blue lines, respectively) obtained from the fitting. The corresponding residuals are plotted at the bottom of each graphic.

**NMR study of the CyHis-EYE complex**

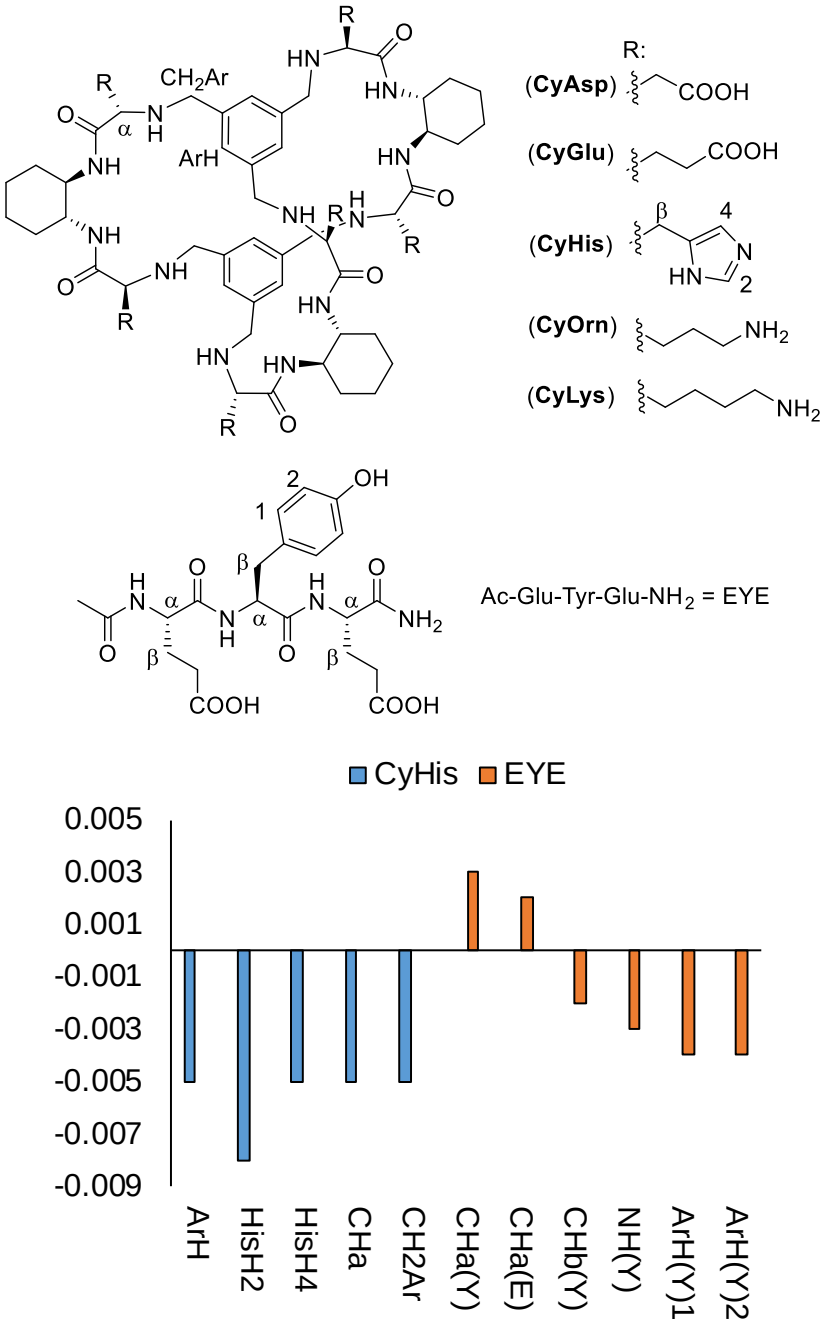


**Figure S9**. Chemical Shift Perturbations (CSP = $\delta_{comp} - \delta_{free}$, ppm) observed by $^{1}$H NMR (500 MHz, 9:1 $H_2O$:$D_2O$ Tris-d11 buffer pH 7.0) for a 1:1 CyHis:EYE sample. Arbitrary signal labelling shown in the corresponding chemical structures.

**Mass Spectrometry characterization of the EYE-Cage complexes**

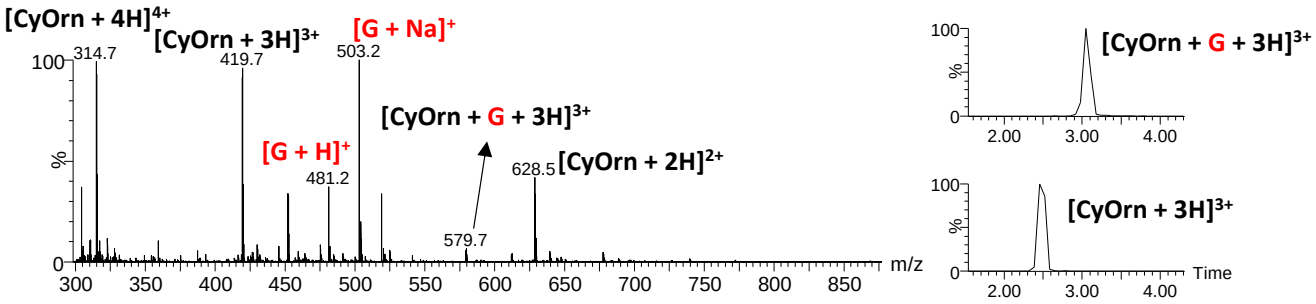


**Figure S10**. ESI TWIM mass spectrum of aqueous 10 μM solution of the **CyOrn** in the presence of 1 equivalent of the tripeptide EYE (left); arrival time distribution for the $[CyOrn + 3H]^{3+}$ and $[CyOrn + G + 3H]^{3+}$ species (right).

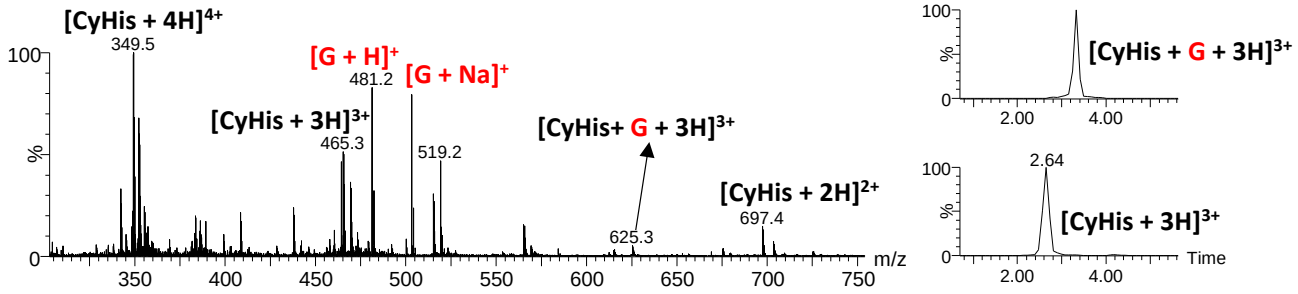


**Figure S11**. ESI TWIM mass spectrum of aqueous 10 μM solution of the **CyHis** in the presence of 1 equivalent of the tripeptide EYE (left); arrival time distribution for the $[CyHis + 3H]^{3+}$ and $[CyHis + G + 3H]^{3+}$ species (right).

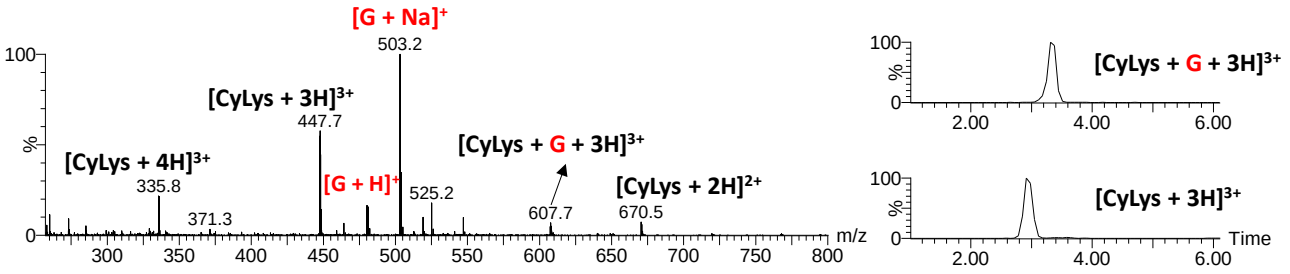


**Figure S12**. ESI TWIM mass spectrum of aqueous 10 μM solution of the **CyLys** in the presence of 3 equivalents of the tripeptide EYE (left); arrival time distribution for the $[CyLys + 3H]^{3+}$ and $[CyLys + G + 3H]^{3+}$ species (right).

**Plot of corrected CCS´ versus effective drift times ($t_D$´) for IM calibrants**

Calibration of the IMS device for determining collision cross-sectional areas from drift time measurements was performed using a mixture of polyalanine reference ions (a series of doubly-charged $[(Ala)n + 2H]^{2+}$ and triply-charged $[(Ala)n + 3H]^{3+}$ species) and Angiotensin I [(Angiotensin I + 3H]$^{3+}$ species covering the transit time of the investigated ions. Their $^{DT}CCS_{N2}$ values were taken from the literature (see refs. 27 and 70 in main text). Table S1 and S2 collect the list of doubly- and triply-charged IM calibrants, respectively, including charge state, experimental drift time, corrected drift time and published collision cross sections $^{DT}\Omega_{N2}$ (Å$^2$)

**Table S1**

| **Sequence** | **$t_D$ (ms)** | **$t_D$' (ms)** | **Published $^{DT}\Omega_{N2}$(Å$^2$)** |
|---|---|---|---|
| $[Ala_{13} + 2H]^{2+}$ | 2.98 | 2.05 | **320** |
| $[Ala_{14} + 2H]^{2+}$ | 3.15 | 2.22 | **333** |
| $[Ala_{15} + 2H]^{2+}$ | 3.38 | 2.45 | **344** |
| $[Ala_{16} + 2H]^{2+}$ | 3.58 | 2.64 | **357** |
| $[Ala_{17} + 2H]^{2+}$ | 3.78 | 2.84 | **369** |
| $[Ala_{18} + 2H]^{2+}$ | 4.01 | 3.07 | **380** |
| $[Ala_{19} + 2H]^{2+}$ | 4.24 | 3.30 | **393** |
| $[Ala_{20} + 2H]^{2+}$ | 4.44 | 3.50 | **404** |
| $[Ala_{21} + 2H]^{2+}$ | 4.64 | 3.70 | **416** |
| $[Ala_{22} + 2H]^{2+}$ | 4.84 | 3.90 | **428** |
| $[Ala_{23} + 2H]^{2+}$ | 5.10 | 4.16 | **437** |
| $[Ala_{24} + 2H]^{2+}$ | 5.37 | 4.43 | **448** |
| $[Ala_{25} + 2H]^{2+}$ | 5.63 | 4.69 | **458** |

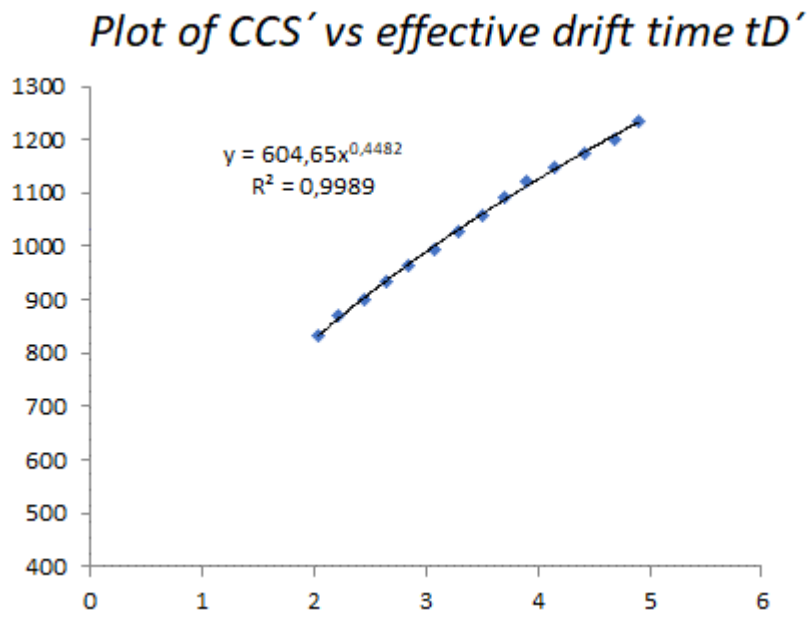


**Figure S12.** Plot of corrected cross sections CCS´ versus effective drift times ($t_D$´) for doubly-charged calibrants. The literature CCS values were converted to CCS´ according to $CCS' = \frac{CCS\sqrt{\mu}}{z}$ where μ and z stands for the reduced mass of the collision partners and the charge state, respectively. A line of best fit plotted through these points is used to obtain collision cross sections from the drift times measured for the supramolecular complexes investigated in this work.

**Table S2**

| Sequence | $t_D$ (ms) | $t_D'$ (ms) | Published $^{DT}\Omega_{N2}$(Å$^2$) |
|---|---|---|---|
| [Angiotensin I + 3H]3+ | 2.94 | 2.01 | **475** |
| $[Ala_{22} + 3H]^{3+}$ | 3.51 | 2.58 | **518** |
| $[Ala_{23} + 3H]^{3+}$ | 2.58 | 2.64 | **532** |
| $[Ala_{24} + 3H]^{3+}$ | 3.78 | 2.84 | **545** |
| $[Ala_{25} + 3H]^{3+}$ | 4.11 | 3.17 | **561** |
| $[Ala_{26} + 3H]^{3+}$ | 4.31 | 3.37 | **576** |
| $[Ala_{27} + 3H]^{3+}$ | 4.31 | 3.37 | **592** |
| $[Ala_{28} + 3H]^{3+}$ | 4.44 | 3.50 | **606** |

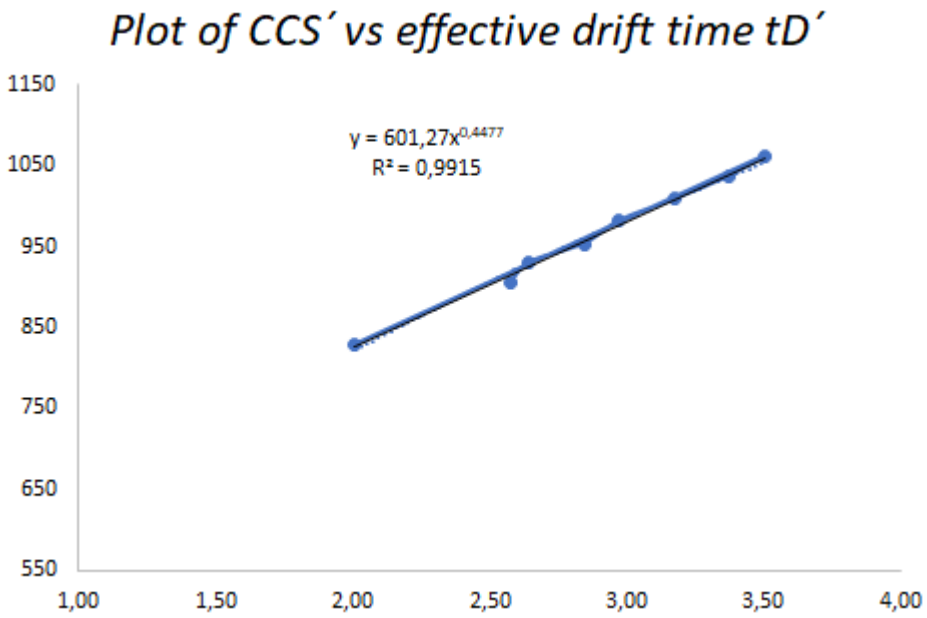


**Figure S13.** Plot of corrected cross sections CCS´ versus effective drift times ($t_D$´) for triply-charged calibrants. The literature CCS values were converted to CCS´ according to $CCS' = \frac{CCS\sqrt{\mu}}{z}$ where μ and z stands for the reduced mass of the collision partners and the charge state, respectively. A line of best fit plotted through these points is used to obtain collision cross sections from the drift times measured for the supramolecular complexes investigated in this work.

**Table S3**. Drift times measured for the doubly- and triply charged species and cross-sectional area extracted by comparison with a series of calibrants.

| **Host:guest complex** | **Charge** | **m/z** | **$t_D$ (ms)** | **$t_D'$ (ms)** | **Experimental $^{TW}CCS_{N2}$ (Å$^2$)** |
|---|---|---|---|---|---|
| [**CyOrn + 3H**]$^{3+}$ | 3 | 419.7 | 2.44 | 1.51 | 414 |
| [**CyOrn + 2H**]$^{2+}$ | 2 | 628.5 | 4.35 | 341 | 401 |
| [**CyOrn + EYE + 3H**]$^{3+}$ | 3 | 579.7 | 3.05 | 2.11 | 480 |
| [**CyLys + 3H**]$^{3+}$ | 3 | 447.4 | 2.91 | 1.98 | 467 |
| [**CyLys + 2H**]$^{2+}$ | 2 | 670.5 | 5.41 | 4.47 | 448 |
| [**CyLys + EYE + 3H**]$^{3+}$ | 3 | 607.3 | 3.39 | 2.46 | 515 |
| [**CyHis + 3H**]$^{3+}$ | 3 | 465.3 | 2.64 | 1.71 | 437 |
| [**CyHis + 2H**]$^{2+}$ | 2 | 697.4 | 4.23 | 3.29 | 395 |
| [**CyHis + EYE + 3H**]$^{3+}$ | 3 | 625.3 | 3.32 | 2.38 | 506 |
| [**CyAsp + 3H**]$^{3+}$ | 3 | 425.8 | 2.41 | 1.48 | 411 |
| [**CyAsp + 2H**]$^{2+}$ | 2 | 638.3 | 3.91 | 2.97 | 377 |
| [**CyGlu + 3H**]$^{3+}$ | 3 | 449.3 | 2.45 | 1.52 | 415 |
| [**CyGlu + 2H**]$^{2+}$ | 2 | 673.4 | 4.09 | 3.15 | 388 |