

Supporting Information for:

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

Ignacio Alfonso,<sup>a,\*</sup> Michael Bolte,<sup>b</sup> Miriam Bru,<sup>c</sup> M. Isabel Burguete,<sup>c</sup> Santiago V. Luis,<sup>c,\*</sup> Cristian Vicent<sup>d</sup>

<sup>a</sup>*Departamento de Química Biológica y Modelización Molecular, Instituto de Química Avanzada de Cataluña, Consejo Superior de Investigaciones Científicas, IQAC / CSIC, C/ Jordi Girona 18-26, E-08034, Barcelona, Spain, Fax: +34-932-045-904; Tel: +34-934-006-158; E-mail: iarqob@iiqab.csic.es*

<sup>b</sup>*Institut für Anorganische Chemie, J. W. Goethe-Universität, Frankfurt, Max-von-Laue-Str. 7, 60438, Frankfurt/Main, Germany.*

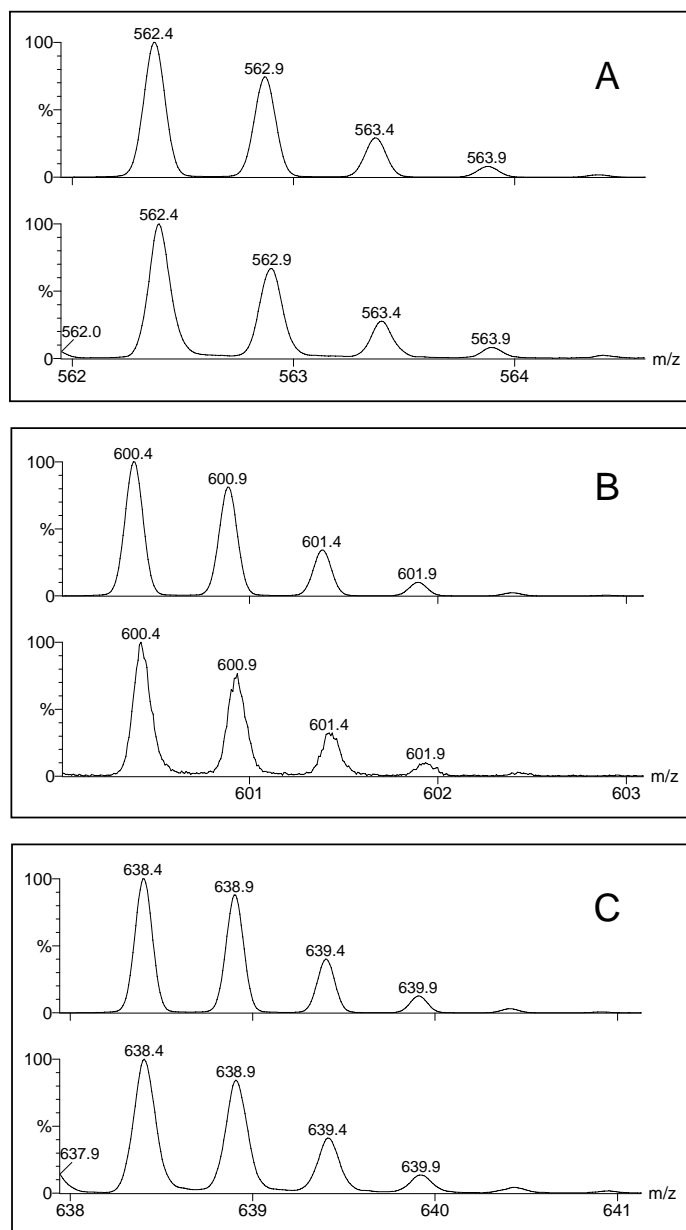
<sup>c</sup>*Departamento de Química Inorgánica y Orgánica/UAMOA, Universidad Jaume I/CSIC, Avenida Sos Baynat s/n, E-12071, Castellón, Spain. Fax: +34-964-728-214; Tel: +34-964-728-238; E-mail: luiss@qio.uji.es*

<sup>d</sup>*Serveis Centrals d'Instrumentació Científica, Universitat Jaume I, Castellón, Spain*

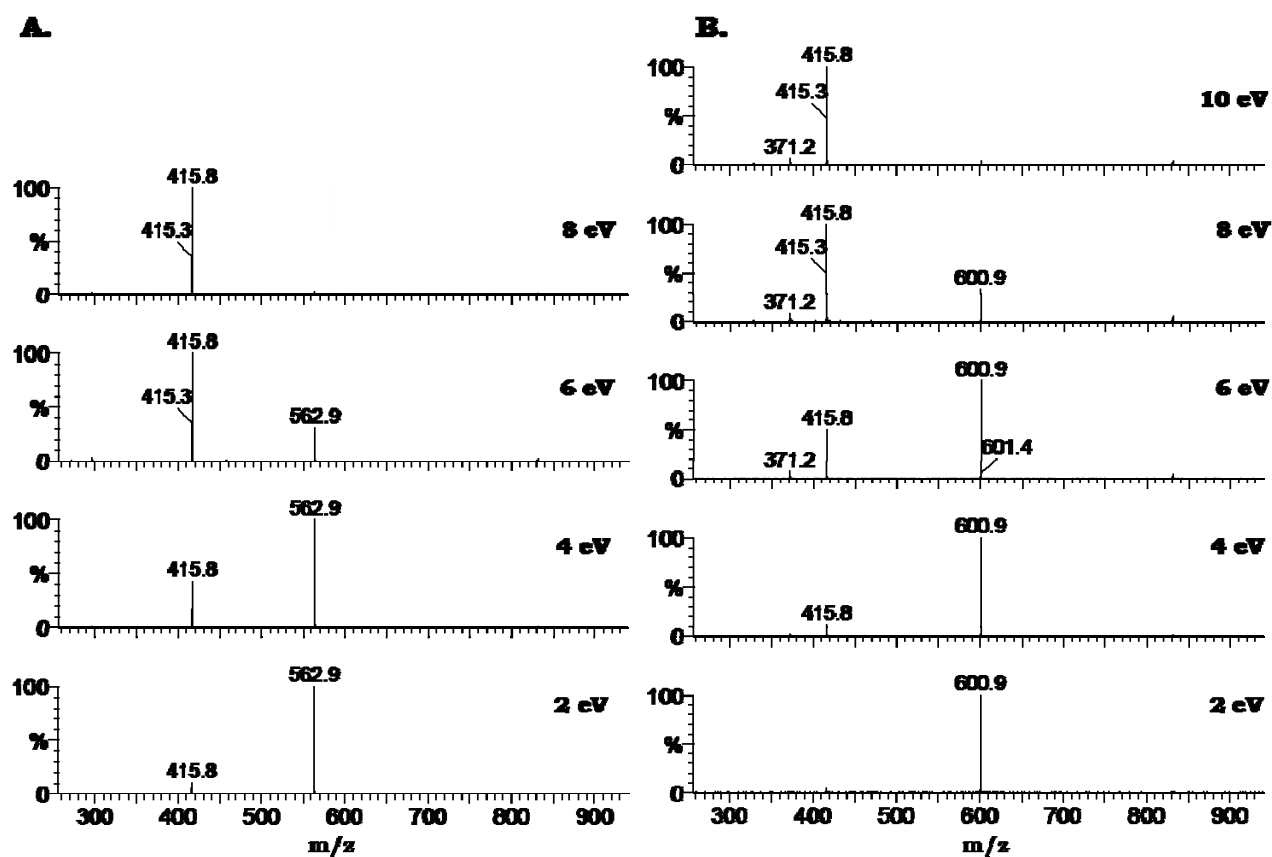
**Table of contents:**

|   |     |
|---|-----|
| High resolution ESI-TOF MS spectra (Figure S1)                              | S2  |
| CID spectra (Figure S2)   | S3  |
| NMR spectra of dipeptides (Figures S3-S9)                                   | S4  |
| NMR spectra of <b>2a</b> (Figures S10-S14)                                  | S8  |
| NOESY/ROESY spectra of the [ <b>2a</b> · Z-AF-OH] complex (Figures S15-S17) | S10 |
| <sup>1</sup> H NMR Titration procedures (Figures S18-S29)                   | S12 |
| X ray diffraction data of the crystals of Z-AF-OH (Table S1)                | S20 |
| Cartesian coordinates for the minimum structure of [ <b>2a</b> · Z-AF-OH]   | S21 |
| Output file from the HYDRONMR program                                       | S22 |

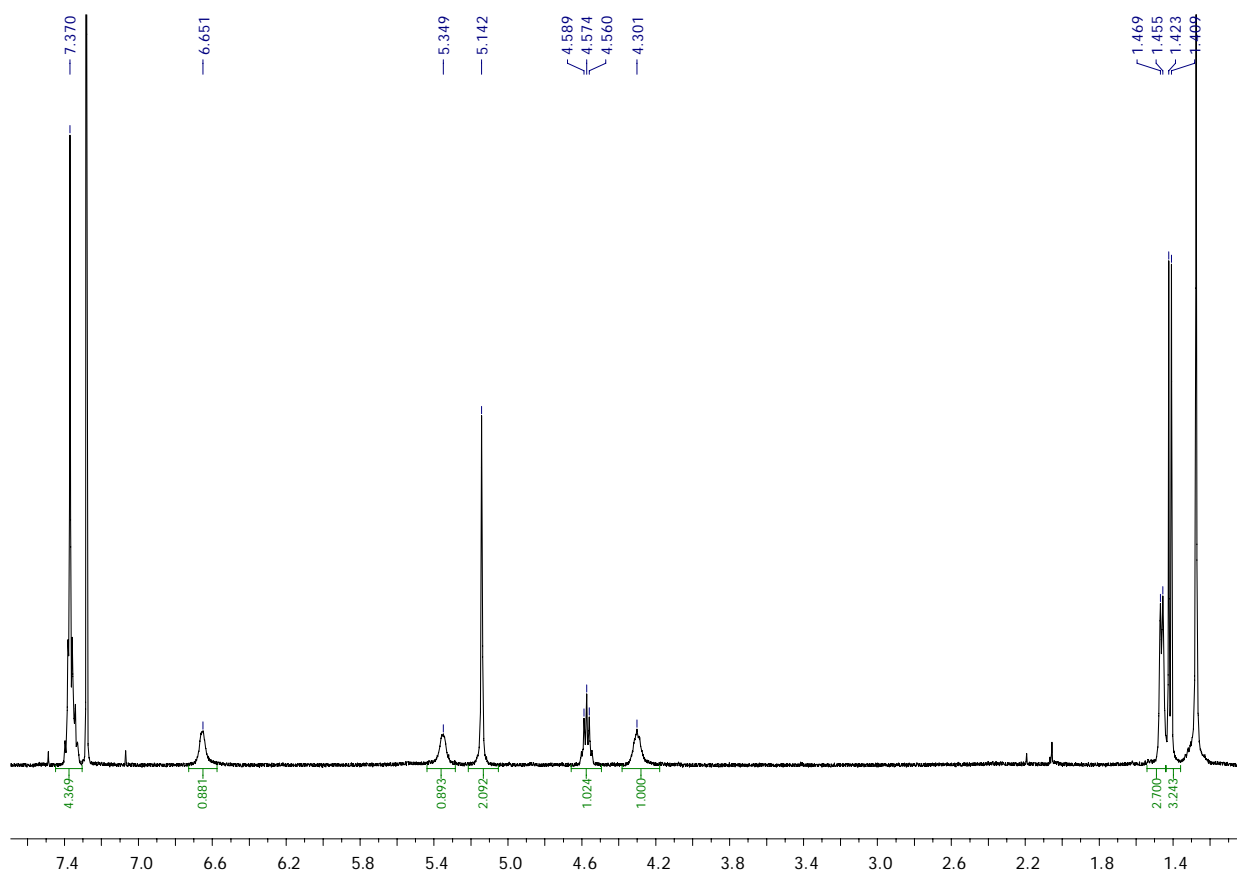
**Figure S1.** Detail of the ESI-TOF spectra for the doubly charged supramolecular complexes formed between **2a** and (a) Z-AA-OH, (b) Z-AF-OH and (c) Z-FF-OH. For every case, the upper and lower traces show the simulated and experimental peaks, respectively.



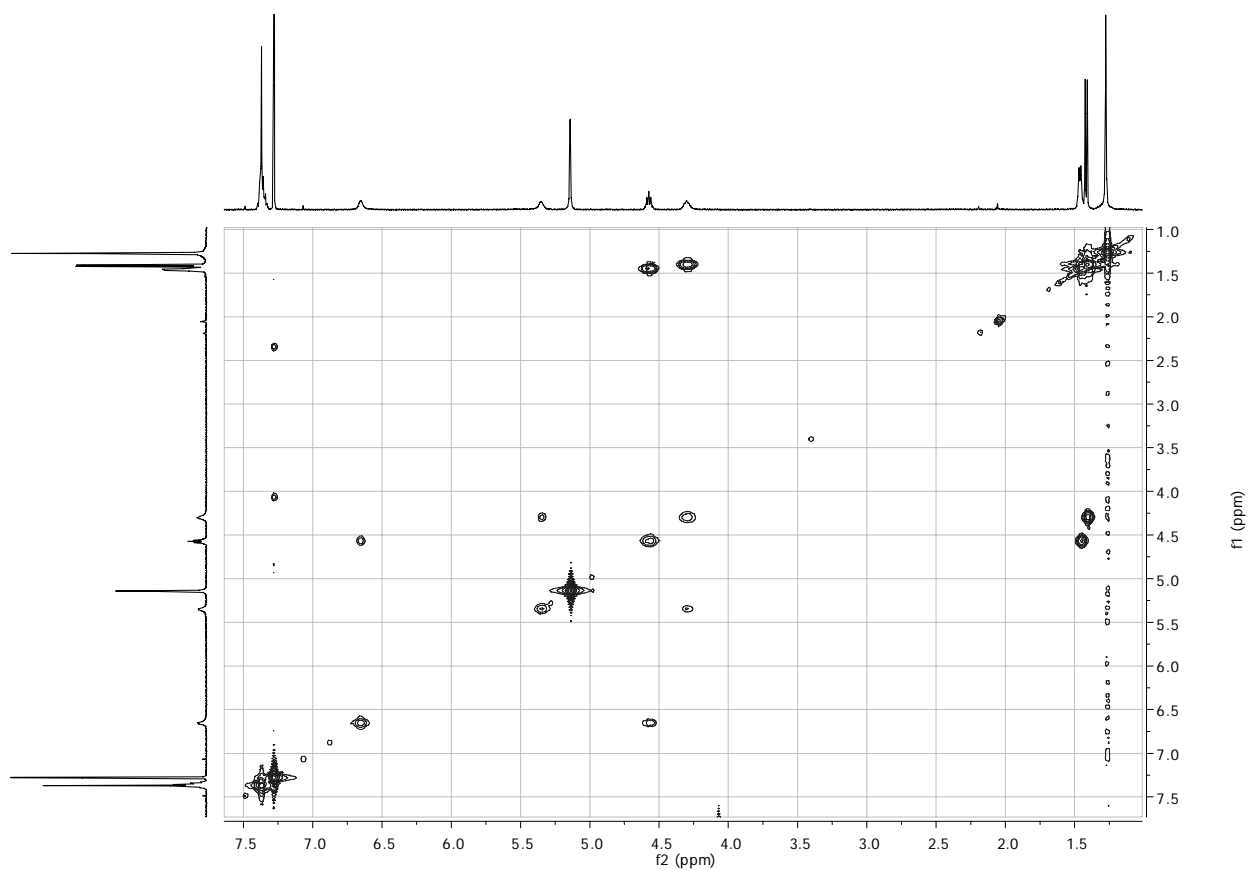
**Figure S2.** Collision Induced Dissociation (CID) spectra of the mass-selected (A)  $[2\mathbf{a} + \text{Z-AA-OH} + 2\text{H}]^{2+}$  and (B)  $[2\mathbf{a} + \text{Z-AF-OH} + 2\text{H}]^{2+}$  supramolecular complexes at increasing collision energies in the  $E_{\text{laboratory}} = 2\text{-}12$  eV range.



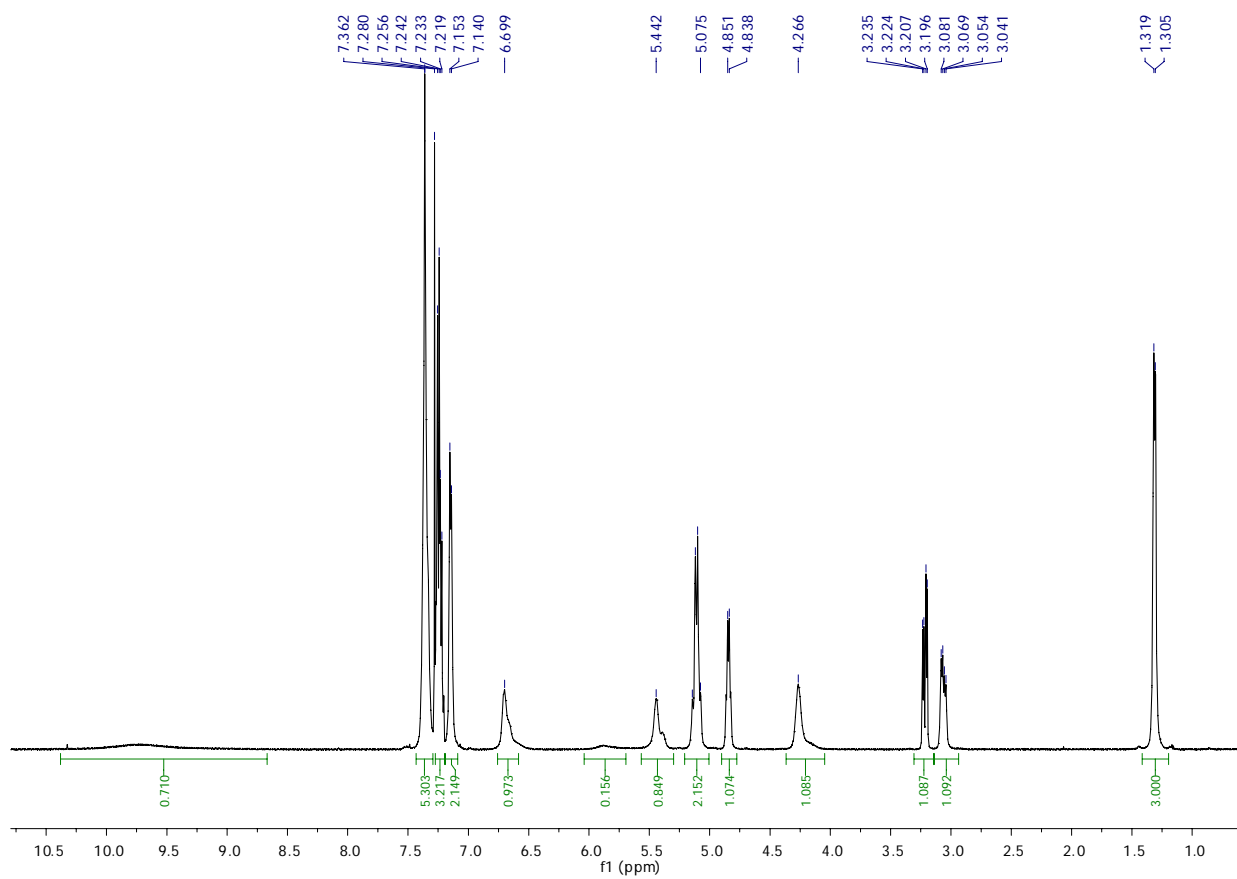
**Figure S3.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-AA-OH.



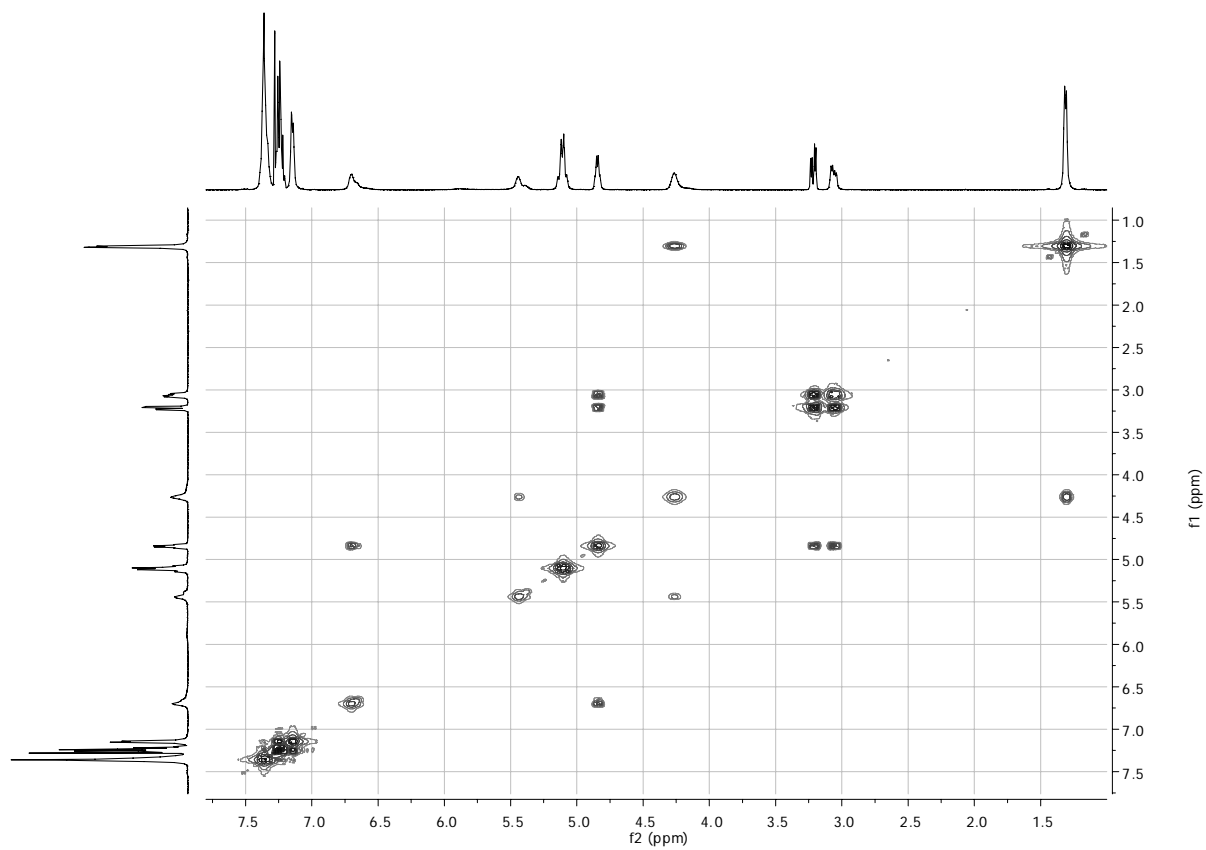
**Figure S4.** gCOSY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-AA-OH.



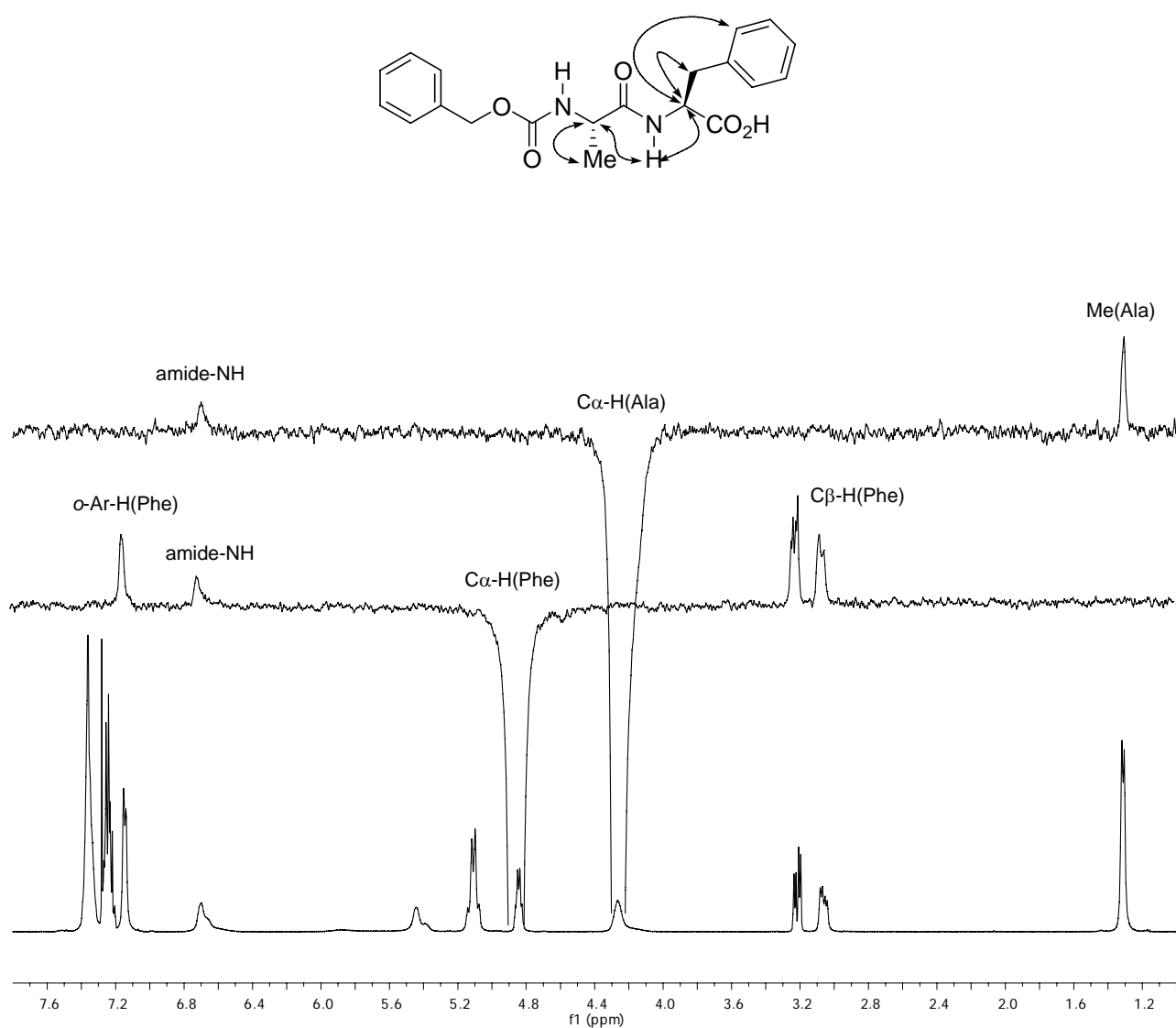
**Figure S5.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-AF-OH.



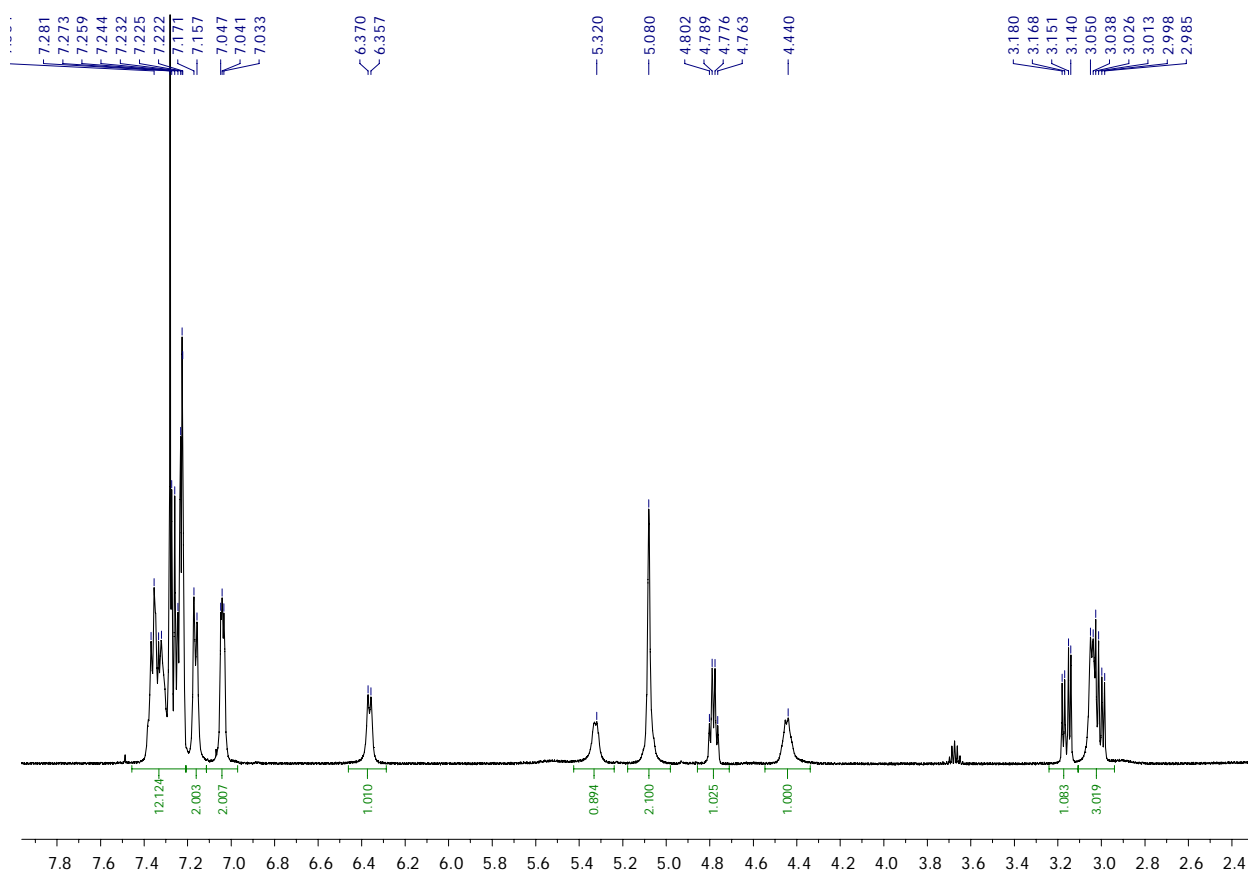
**Figure S6.** gCOSY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-AF-OH.



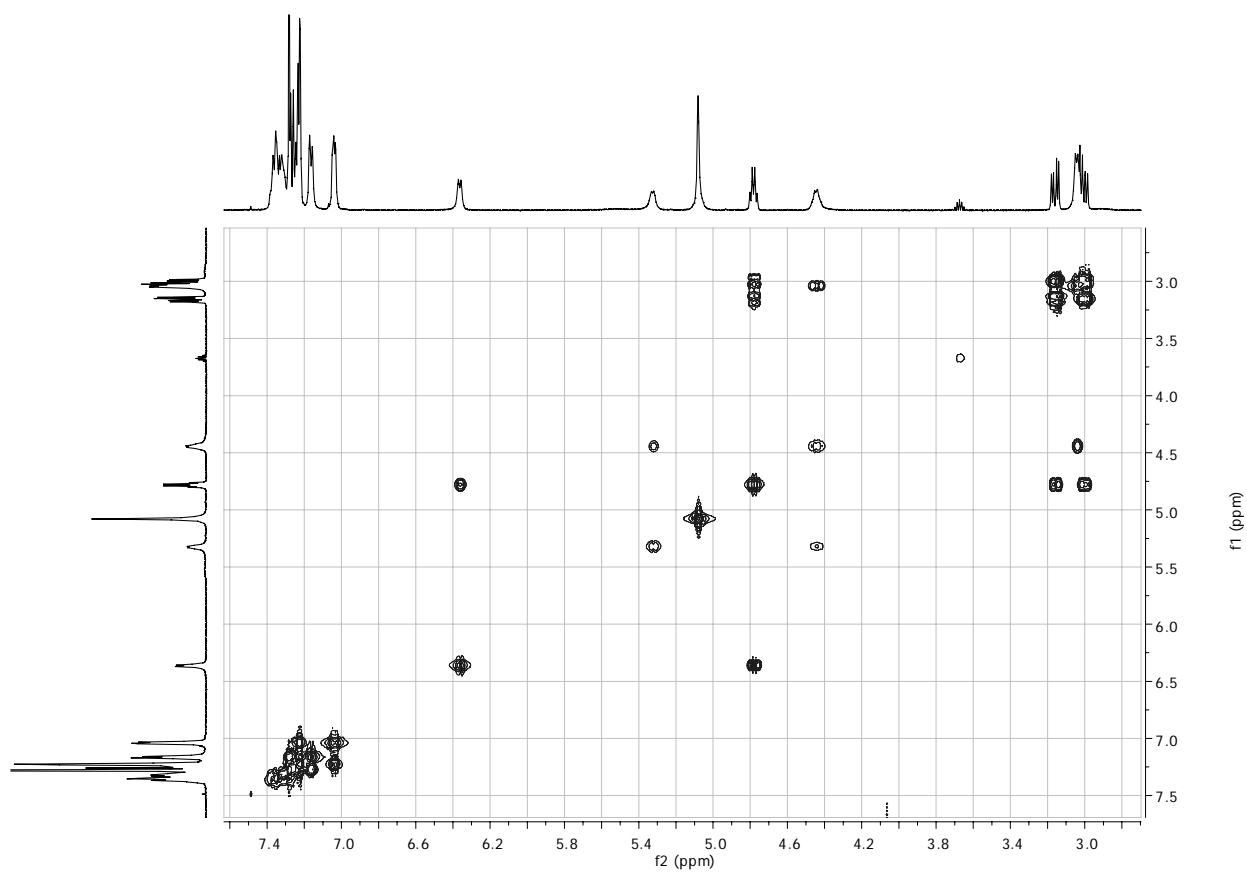
**Figure S7.**  $^1\text{H}$  NMR (lower trace) and 1D NOESY spectra (500 MHz,  $\text{CDCl}_3$ ) of Z-AF-OH upon irradiation of Ala (upper trace) and Phe (middle trace)  $\text{C}\alpha$  protons. The concentration of the sample was 30 mM, and positive nOEs were clearly observed (in antiphase with respect to the corresponding irradiated signal).



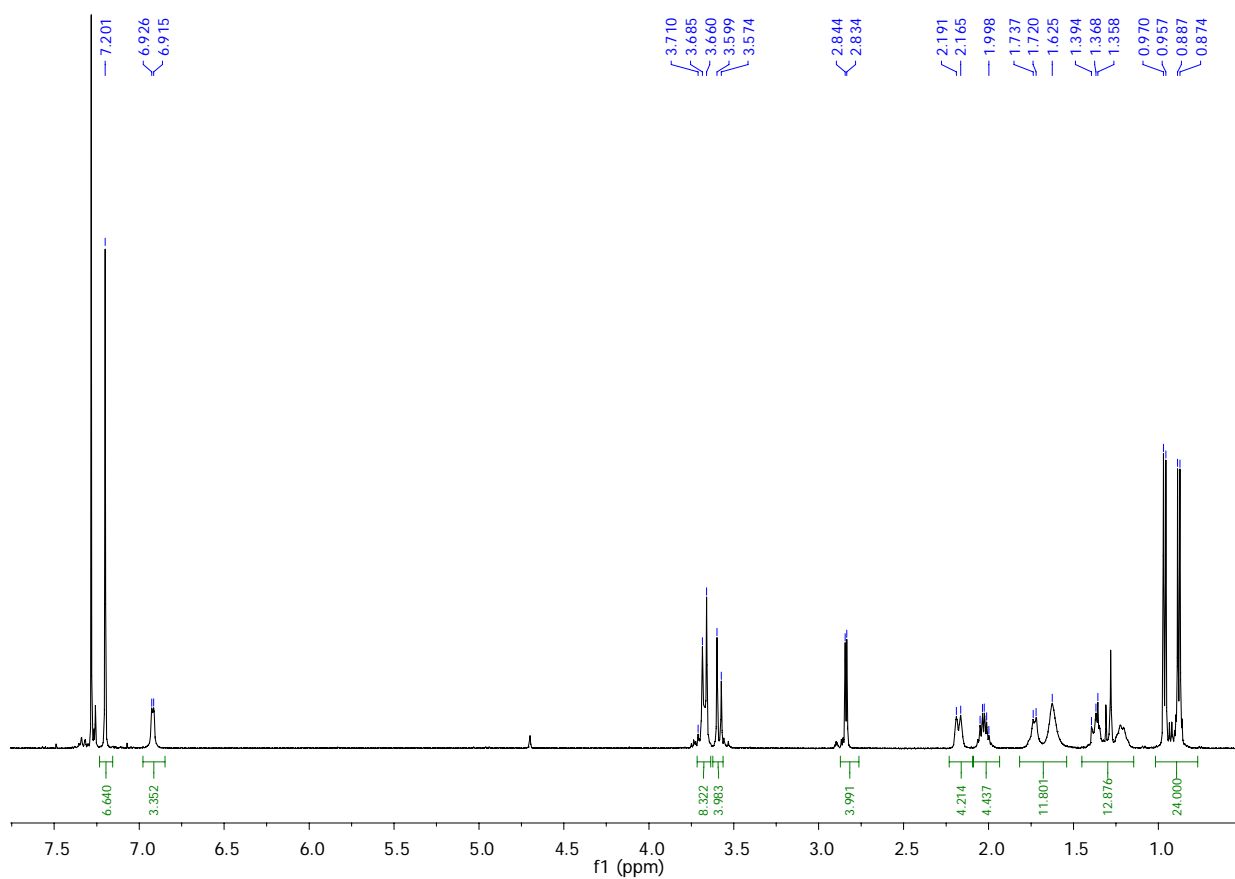
**Figure S8.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-FF-OH.



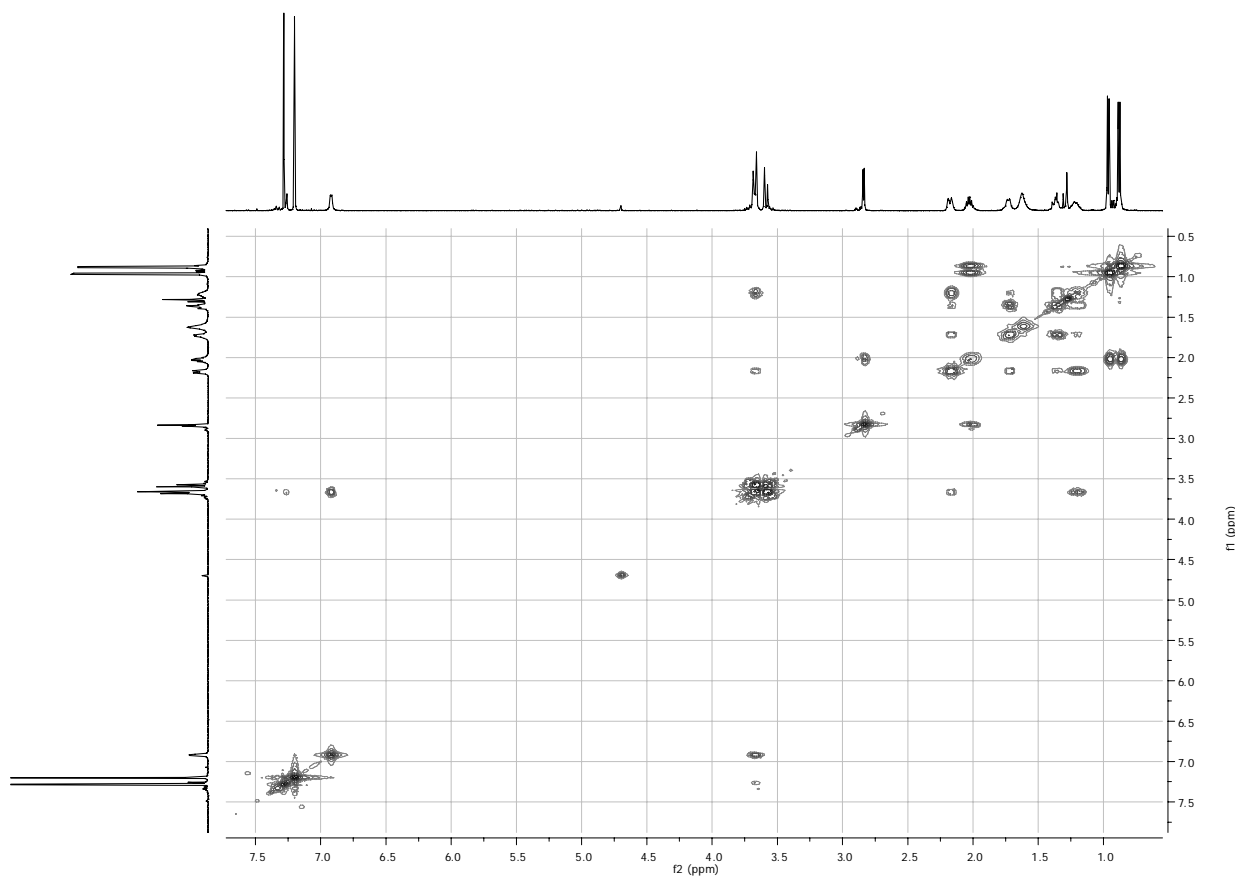
**Figure S9.** gCOSY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of Z-FF-OH.



**Figure S10.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of **2a**.

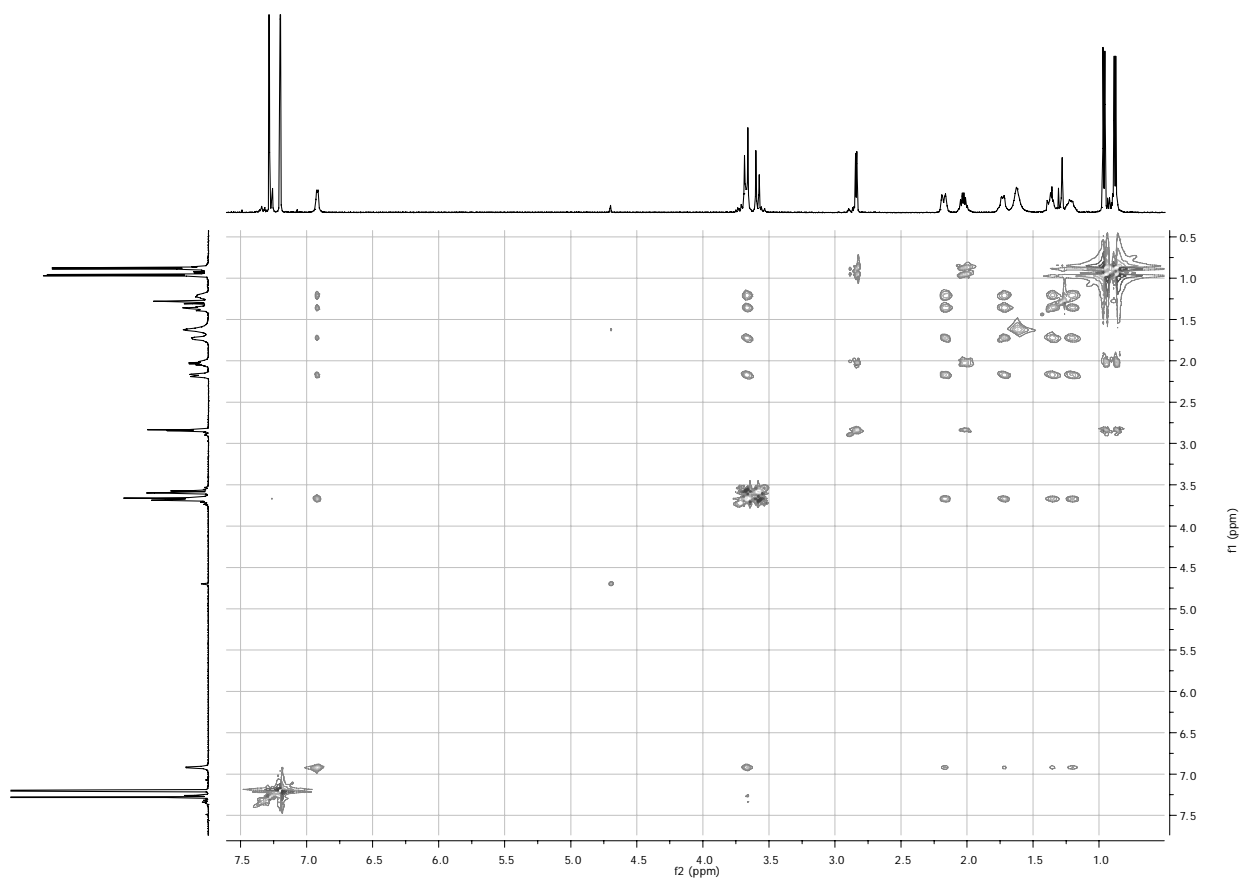


**Figure S11.** gCOSY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of **2a**.

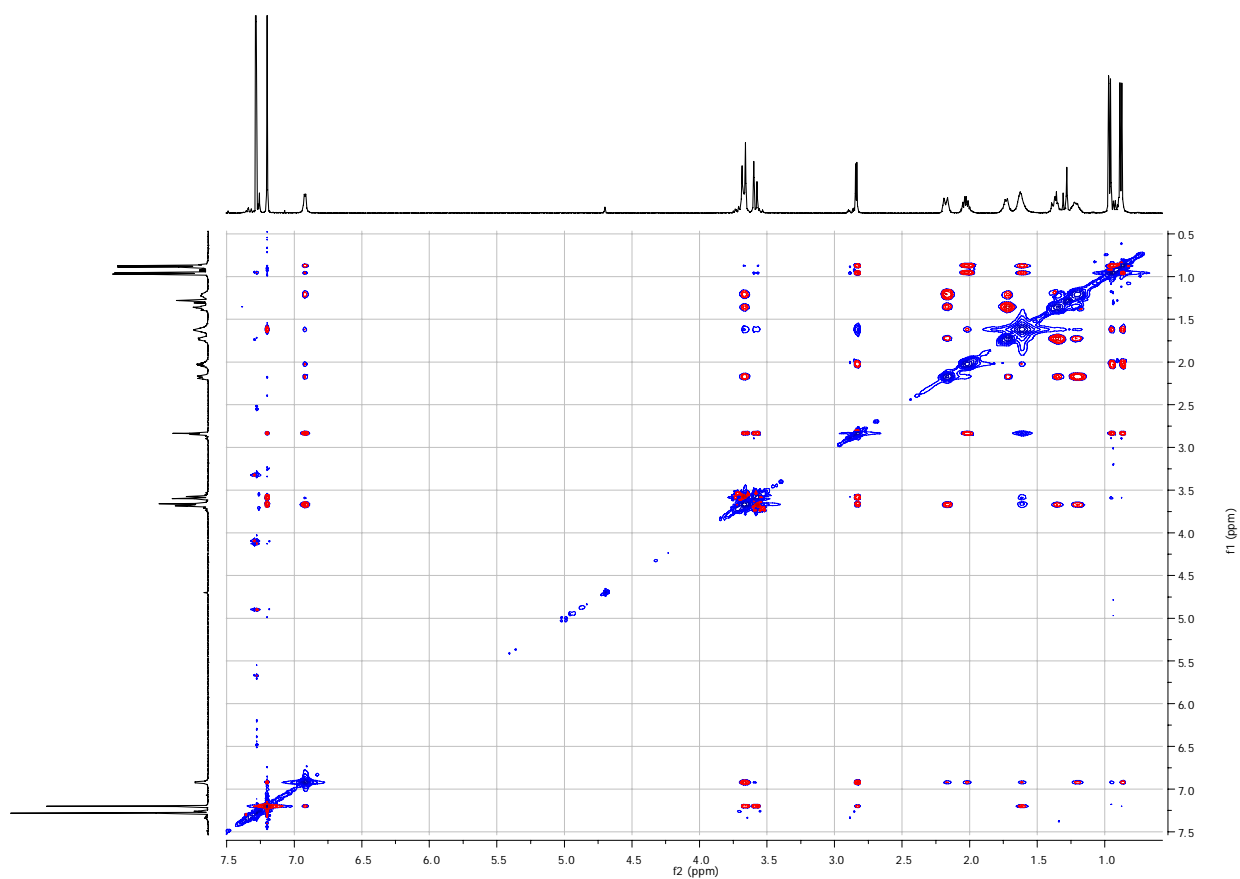




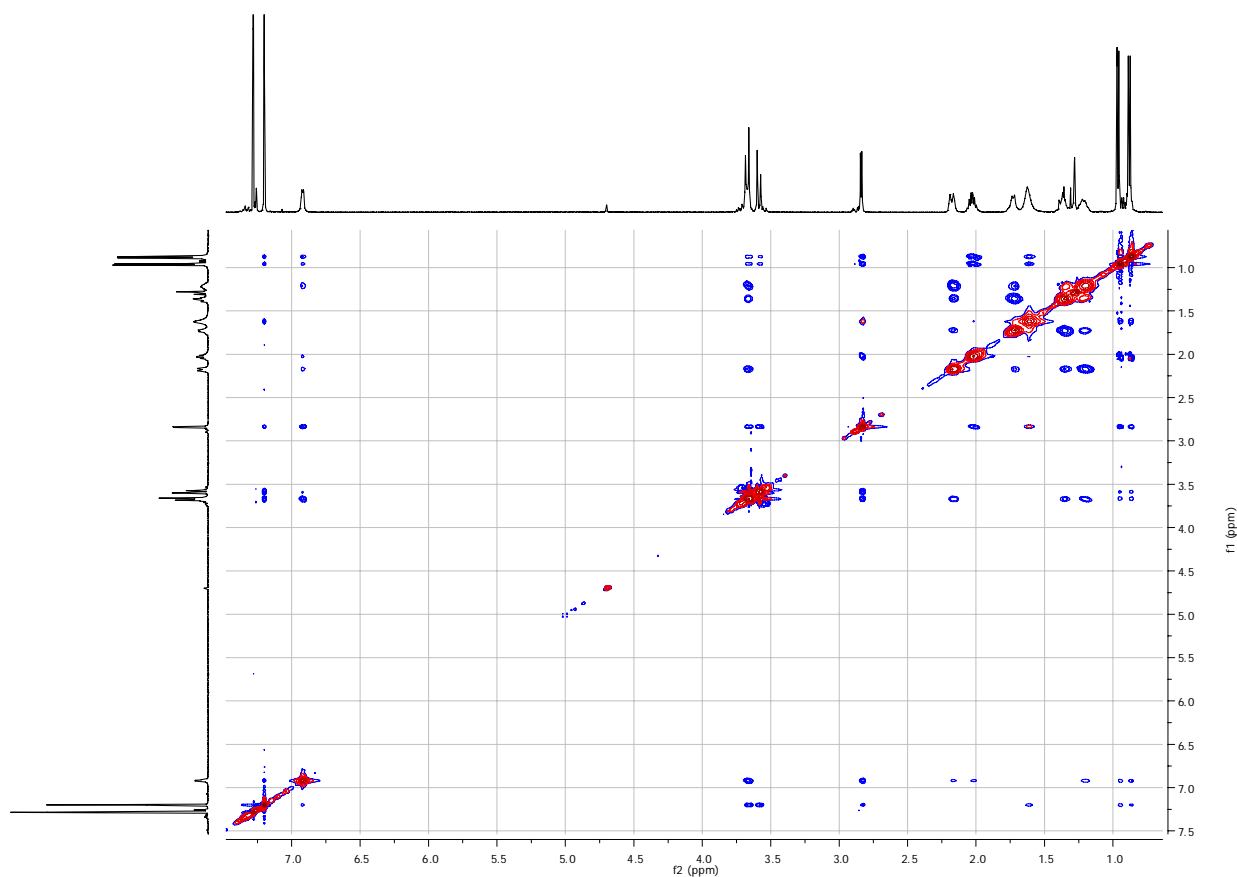
**Figure S12.** gTOCSY NMR spectrum (500 MHz, CDCl<sub>3</sub>) of **2a**.



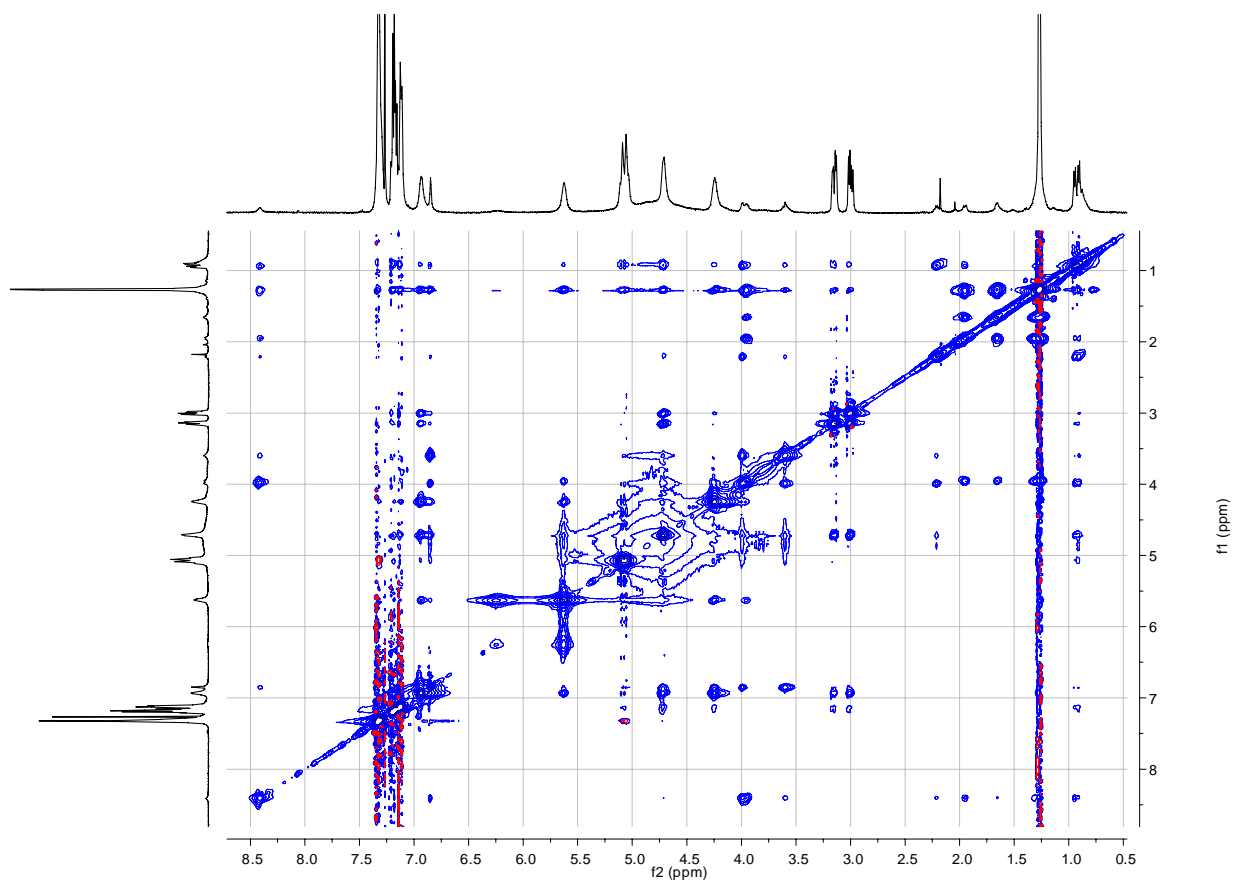
**Figure S13.** NOESY NMR spectrum (500 MHz, CDCl<sub>3</sub>) of **2a**.



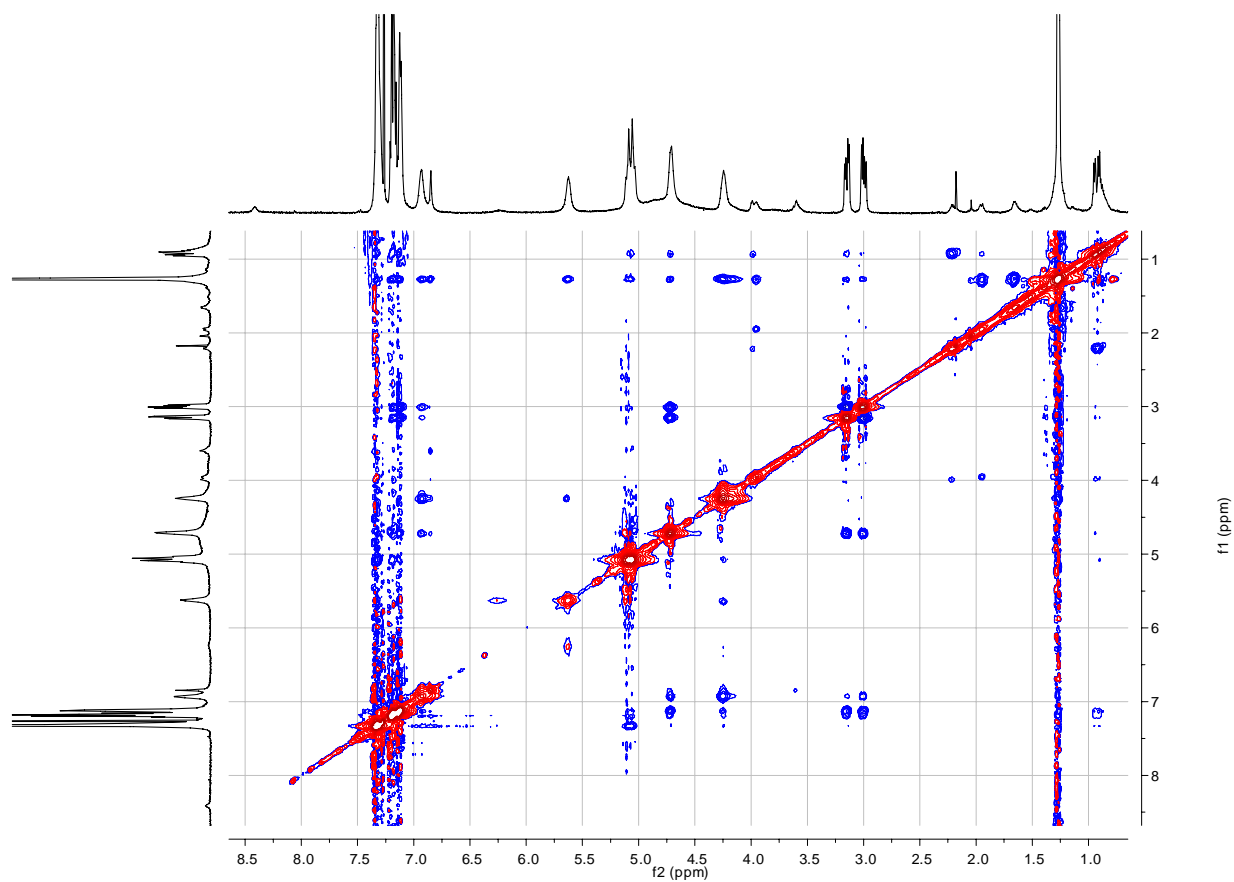
**Figure S14.** ROESY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of **2a**.



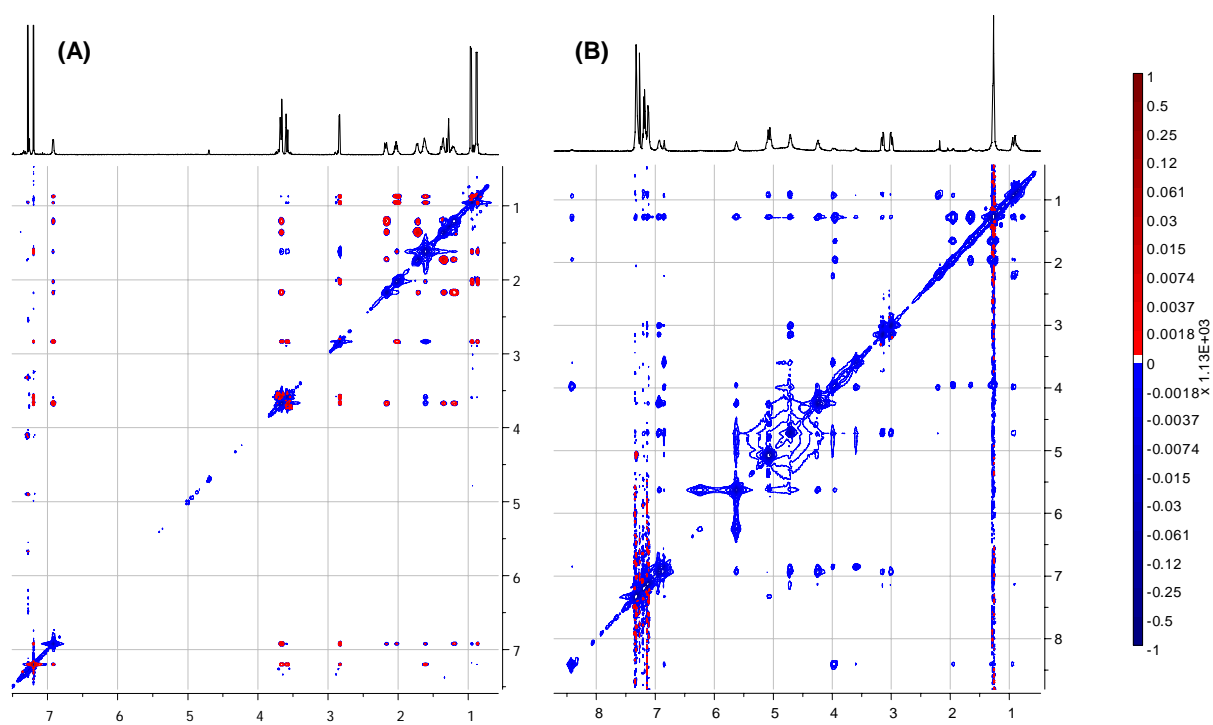
**Figure S15.** NOESY NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of **2a** saturated with Z-AF-OH



**Figure S16.** ROESY NMR spectrum (500 MHz, CDCl<sub>3</sub>) of **2a** saturated with Z-AF-OH



**Figure S17.** Comparison of the NOESY NMR spectra (500 MHz, CDCl<sub>3</sub>) of **2a** alone (A) and saturated with Z-AF-OH (B)



## <sup>1</sup>H NMR Titration procedures

All the NMR titrations were performed at 303 K in a Varian INOVA 500 spectrometer operating at 500 MHz for proton. The non-linear least squares fitting calculations were performed with KaleidaGraph versio 4.03 by Synergy Software.

**Titration procedure for Z-F-OH:** We prepared a 3.3 mM stock solution of **2a** in dry, acid free CDCl<sub>3</sub> containing 1% MeOH. For the titrant solutions, weighted amounts of Z-F-OH were dissolved in the stock solution of the receptor (0.5-1.0 M final concentration of Z-F-OH) and thus, the solution of the titrant also contained a 3.3 mM concentration of **2a** in order to keep the concentration of the receptor constant during the titration experiment. Then, 0.5 mL of the stock solution of the receptor was placed in a NMR tube and the <sup>1</sup>H NMR spectrum was acquired. Afterwards, the receptor was carefully titrated by addition of small amounts of titrant solutions of the substrate. The <sup>1</sup>H NMR spectra were acquired (500 MHz, 303 K) after each addition, using 16-32 scans and a relaxation delay of 5 s. Different <sup>1</sup>H NMR signals of **2a** were monitored during the experiment, although the most suitable for quantitative analysis was that corresponding to the receptor amide NH protons.

We used the simplest 1 : 1 binding model to fit the data, using the non-linear least squares fitting of the experimental values to the following equation:<sup>1</sup>

$$\Delta\delta = \frac{\Delta\delta_{\max} K_{\text{as}} [\text{L}]_{\text{tot}}}{1 + K_{\text{as}} [\text{L}]_{\text{tot}}}$$

Where:  $\Delta\delta$  (y, dependent variable) is the observed variation of the chemical shift ( $\delta_{\text{obs}} - \delta_{\text{o}}$ ) being

$\delta_{\text{obs}}$  the chemical shift at each titration point and  $\delta_{\text{o}}$  the initial chemical shift (host alone)

$[\text{L}]_{\text{tot}}$  is the overall concentration of the guest titrant (m0, independent variable)

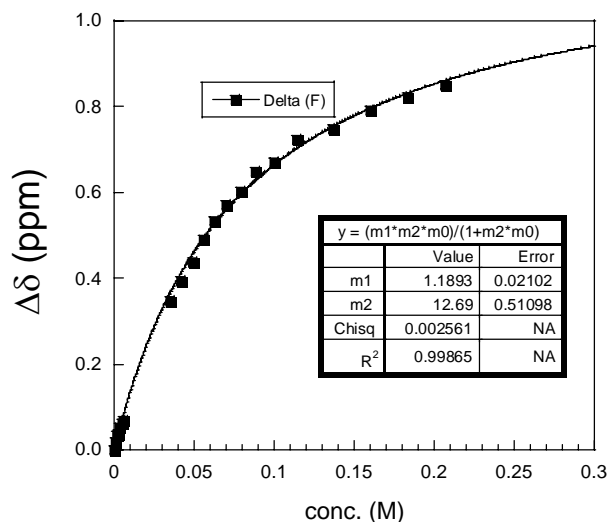
$\Delta\delta_{\max}$  is the maximum variation of the observed chemical shift (m1 in the fitting)

$K_{\text{as}}$  is the host-guest binding constant (m2 in the fitting)

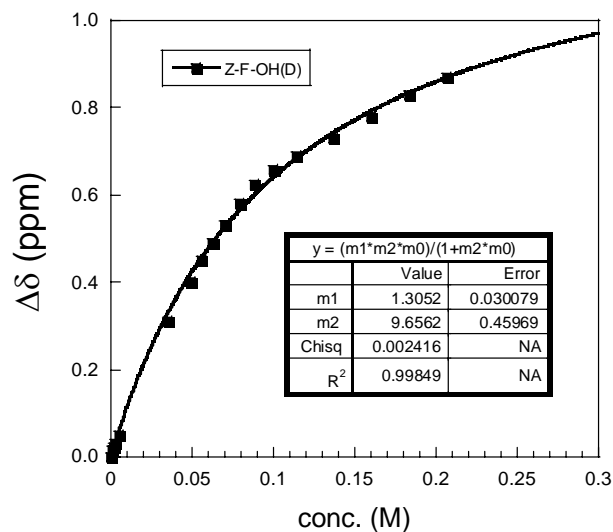
---

<sup>1</sup> Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability*, John Wiley & Sons, 1987.

**Figure S18.** Titration of Z-F-OH(L) (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>): the symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



**Figure S19.** Titration of Z-F-OH(D) (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>): the symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



**Self-assembling of dipeptides:** Solutions of the dipeptides in CDCl<sub>3</sub> containing 1% of MeOH were prepared at different final concentrations (0.1-90 mM, depending on peptide solubility). The <sup>1</sup>H NMR spectra were acquired (500 MHz, 303 K) for every sample, using 32-64 scans and a relaxation delay of 5 s. Different <sup>1</sup>H NMR signals were monitored during the experiment, although the most suitable for quantitative analysis were those corresponding to the amide NH protons. We used the simplest monomer-dimer equilibrium model to successfully fit the experimental data using the non-linear least squares fitting to the following equation:<sup>2</sup>

$$\delta_{\text{obs}} = \delta_{\text{m}} + \frac{1 + 4C_{\text{tot}}K_{\text{dim}} - \sqrt{1 + 8C_{\text{tot}}K_{\text{dim}}}}{4C_{\text{tot}}K_{\text{dim}}} \times (\delta_{\text{d}} - \delta_{\text{m}})$$

Where:  $\delta_{\text{obs}}$  is the observed chemical shift at each titration point (y, dependent variable)

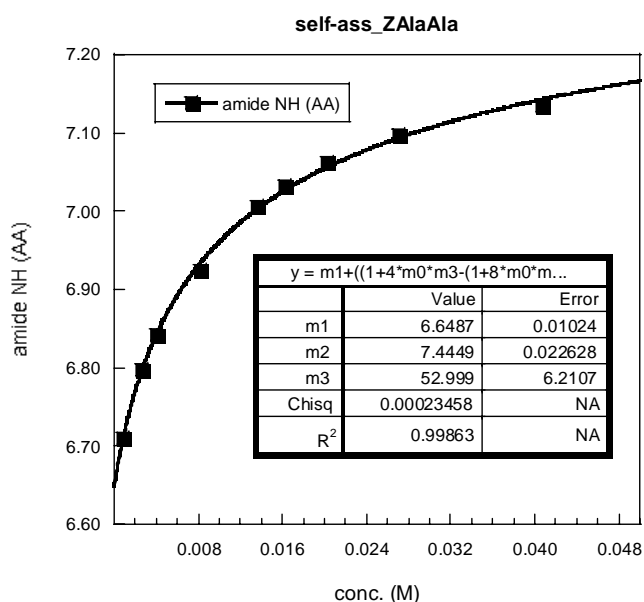
$C_{\text{tot}}$  is the overall concentration of the compound (m0, independent variable)

$\delta_{\text{m}}$  is the chemical shift of the monomer (m1 in the fitting)

$\delta_{\text{d}}$  is the chemical shift of the dimer (m2 in the fitting)

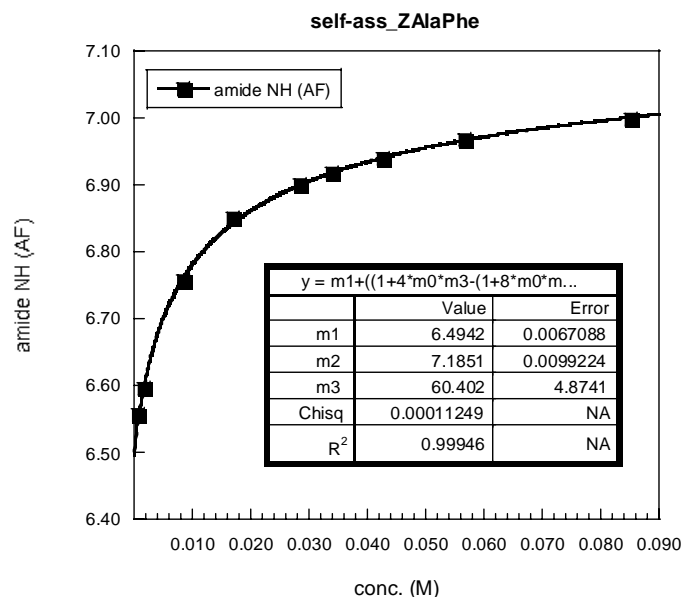
$K_{\text{dim}}$  is the dimerization equilibrium constant (m3 in the fitting)

**Figure S20.** Dilution titration of Z-AA-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>): the symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.

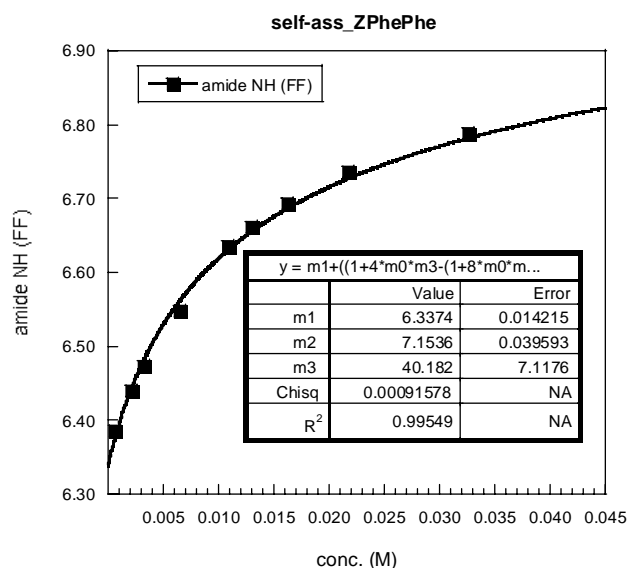


<sup>2</sup> (a) Schmuck, C.; Rehm, T.; Geiger, L.; Schäfer, M. *J. Org. Chem.* **2007**, *72*, 6162. (b) Martin, R. B. *Chem. Rev.* **1996**, *96*, 3043. (c) Davis, J. C. Jr.; Deb, K. K. *Adv. Magn. Reson.* **1970**, *4*, 201. (d) Bangerter, B. W.; Chan, S. I. *J. Am. Chem. Soc.* **1969**, *91*, 3910.

**Figure S21.** Dilution titration of Z-AF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>): the symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



**Figure S22.** Dilution titration of Z-FF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>): the symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



**Host-guest binding by dilution titration:** Samples of exactly equimolar amounts of **2a** or **1a** and a given substrate were prepared in  $\text{CDCl}_3$  containing 1% of MeOH and at different final global concentrations (1-20 mM). The  $^1\text{H}$  NMR spectra were acquired (500 MHz, 303 K) for every sample, using 32-64 scans and a relaxation delay of 5 s. Different  $^1\text{H}$  NMR signals were monitored during the experiment, although the most suitable for quantitative analysis were those corresponding to the dipeptide amide NH protons.

We used the simplest 1 : 1 binding model to fit the data, using the non-linear least squares fitting of the experimental values to the following equation:<sup>3</sup>

$$\delta_{\text{obs}} = \delta_o - \frac{(\delta_o - \delta_c)}{2C_{\text{tot}}K_{\text{as}}} \times \left( 2C_{\text{tot}}K_{\text{as}} + 1 - \sqrt{4C_{\text{tot}}K_{\text{as}} + 1} \right)$$

Where:  $\delta_{\text{obs}}$  is the observed chemical shift at each titration point (y, dependent variable)

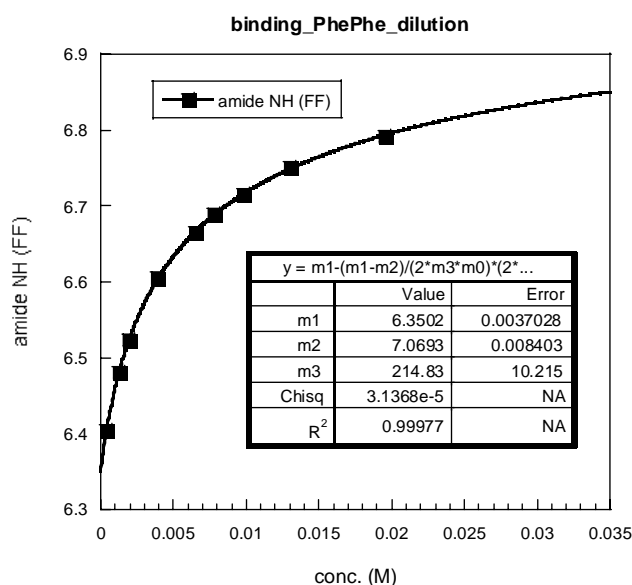
$C_{\text{tot}}$  is the overall concentration of both components (m0, independent variable)

$\delta_o$  is the chemical shift of the unbound species (m1 in the fitting)

$\delta_c$  is the chemical shift of the complex species (m2 in the fitting)

$K_{\text{as}}$  is the host-guest binding constant (m3 in the fitting)

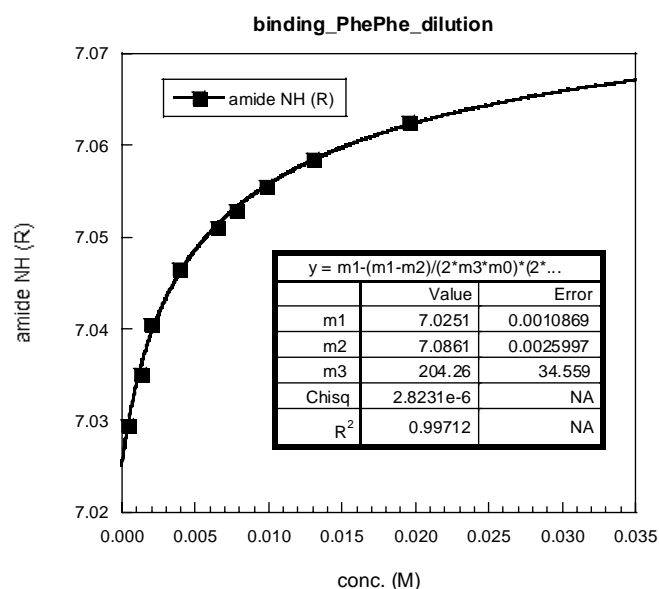
**Figure S23.** Dilution titration of a 1 : 1 mixture of **2a** : Z-FF-OH (500 MHz, 303 K, 1% MeOH in  $\text{CDCl}_3$ ) using the dipeptide amide NH proton signal. The symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



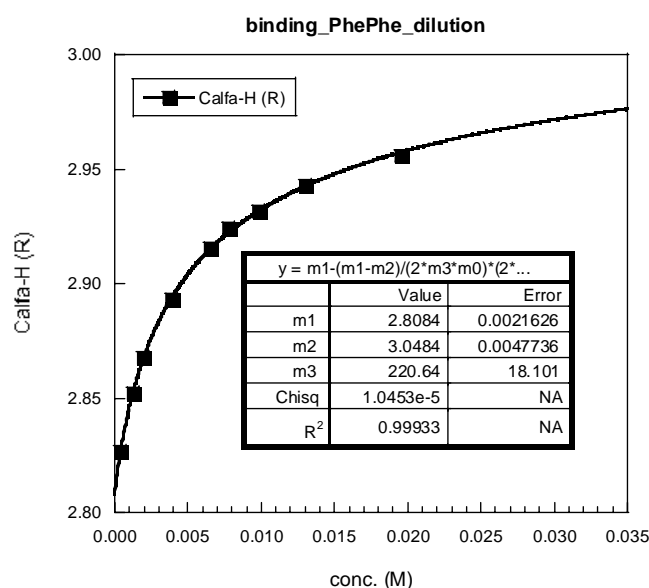
<sup>3</sup> (a) Peters, L.; Fröhlich, R.; Boyd, A. S. F.; Kraft, A. *J. Org. Chem.* **2001**, *66*, 3291. (b) Macomber, R. S. *J. Chem. Educ.* **1992**, *69*, 375. (c) Schneider, H. J.; Kramer, R.; Simova, S.; Schneider, U. *J. Am. Chem. Soc.* **1988**, *110*, 6442.



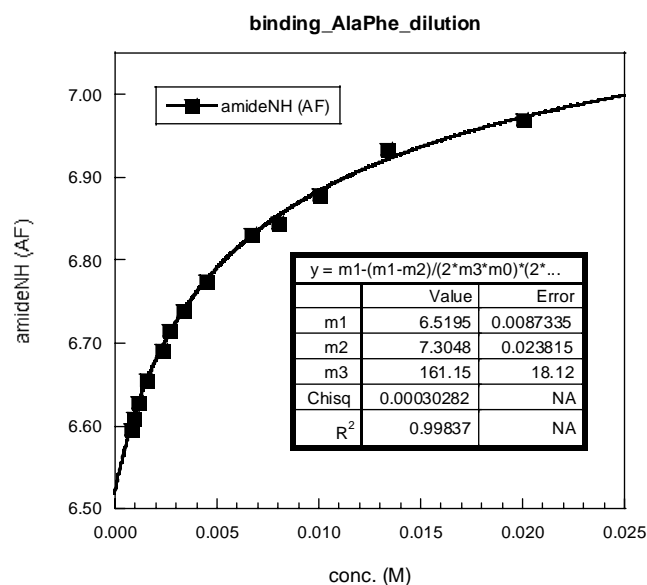
**Figure S24.** Dilution titration of a 1 : 1 mixture of **2a** : Z-FF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the receptor amide NH proton signal. The symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



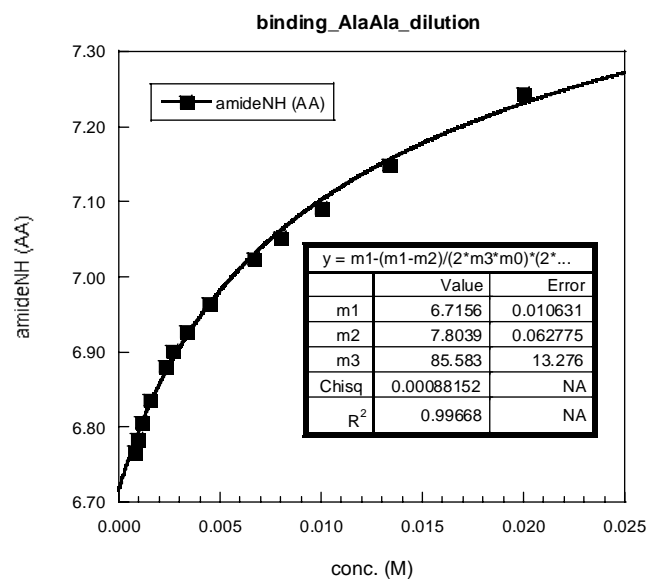
**Figure S25.** Dilution titration of a 1 : 1 mixture of **2a** : Z-FF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the receptor C $\alpha$  proton signal. The symbols correspond to experimental data and the line to the non-linear squares least fitting. The final values for the best fitting are given in the figure.



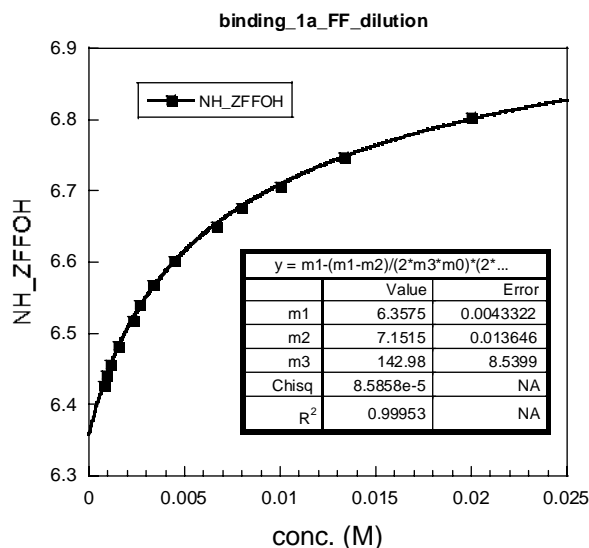
**Figure S26.** Dilution titration of a 1 : 1 mixture of **2a** : Z-AF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the dipeptide amide NH proton signal. The symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.



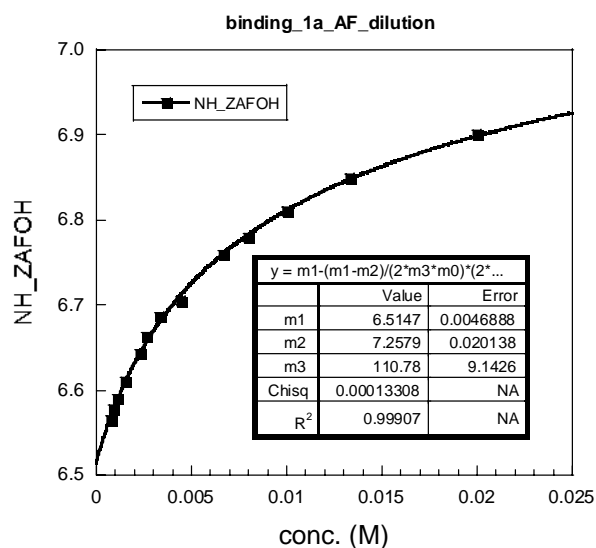
**Figure S27.** Dilution titration of a 1 : 1 mixture of **2a** : Z-AA-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the dipeptide amide NH proton signal. The symbols correspond to experimental data and the line to the non-linear least squares fitting. The final values for the best fitting are given in the figure.

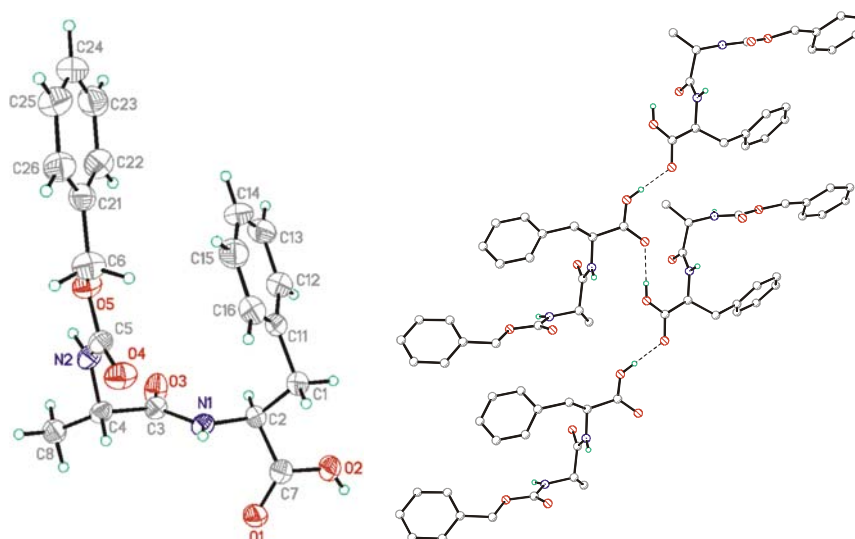


**Figure S28.** Dilution titration of a 1 : 1 mixture of **1a** : Z-FF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the receptor C $\alpha$  proton signal. The symbols correspond to experimental data and the line to the non-linear squares least fitting. The final values for the best fitting are given in the figure.



**Figure S28.** Dilution titration of a 1 : 1 mixture of **1a** : Z-AF-OH (500 MHz, 303 K, 1% MeOH in CDCl<sub>3</sub>) using the receptor C $\alpha$  proton signal. The symbols correspond to experimental data and the line to the non-linear squares least fitting. The final values for the best fitting are given in the figure.





**Table S1. Crystal data and structure refinement for Z-AF-OH.**

|                                   |   |                 |
|-----------------------------------|---|-----------------|
| Identification code               | nacho1  |                 |
| Empirical formula                 | C <sub>20</sub> H <sub>22</sub> N <sub>2</sub> O <sub>5</sub> |                 |
| Formula weight                    | 370.40  |                 |
| Temperature                       | 173(2) K  |                 |
| Wavelength                        | 0.71073 Å   |                 |
| Crystal system                    | Monoclinic  |                 |
| Space group                       | C 2   |                 |
| Unit cell dimensions              | a = 43.977(3) Å   | α = 90°.        |
|                                   | b = 5.0960(3) Å   | β = 95.848(6)°. |
|                                   | c = 8.3187(6) Å   | γ = 90°.        |
| Volume                            | 1854.6(2) Å <sup>3</sup>                                      |                 |
| Z                                 | 4   |                 |
| Density (calculated)              | 1.327 Mg/m <sup>3</sup>                                       |                 |
| Absorption coefficient            | 0.096 mm <sup>-1</sup>  |                 |
| F(000)                            | 784   |                 |
| Crystal size                      | 0.19 x 0.17 x 0.14 mm <sup>3</sup>                            |                 |
| Theta range for data collection   | 1.86 to 25.00°.   |                 |
| Index ranges                      | -48 ≤ h ≤ 52, -5 ≤ k ≤ 6, -9 ≤ l ≤ 9                          |                 |
| Reflections collected             | 7601  |                 |
| Independent reflections           | 1833 [R(int) = 0.0762]  |                 |
| Completeness to theta = 25.00°    | 99.8 %  |                 |
| Absorption correction             | None  |                 |
| Max. and min. transmission        | 0.9867 and 0.9820   |                 |
| Refinement method                 | Full-matrix least-squares on F <sup>2</sup>                   |                 |
| Data / restraints / parameters    | 1833 / 1 / 251  |                 |
| Goodness-of-fit on F <sup>2</sup> | 1.151   |                 |
| Final R indices [I > 2σ(I)]       | R1 = 0.0533, wR2 = 0.1587                                     |                 |
| R indices (all data)              | R1 = 0.0580, wR2 = 0.1607                                     |                 |
| Absolute structure parameter      | 1(3)  |                 |
| Extinction coefficient            | 0.014(3)  |                 |
| Largest diff. peak and hole       | 0.370 and -0.219 e.Å <sup>-3</sup>                            |                 |



Output file from the HYDRONMR program

-----  
HYDRONMR Versi on 7. C

J. García de la Torre, M.L. Huertas and B. Carrasco,  
"HYDRONMR: prediction of NMR relaxation of globular proteins  
from atomic-level structures and hydrodynamic calculations"  
J. Magnetic Reson. 147, 138-146 (2000).

-----  
SUMMARY OF DATA AND RESULTS

This file: comp\_AF1.res  
Case: comp\_AF1  
Structural file: comp\_AF1.pdb  
Temperature: 298.0 K  
Solvent viscosity: 0.00537 poise  
Radius of atomic elements: 2.5 Angs  
Translational diffusion coefficient: 4.543E-06 cm<sup>2</sup>/s  
Center of diffusion (x): 8.049E-08 cm  
Center of diffusion (y): 9.481E-08 cm  
Center of diffusion (z): 1.295E-07 cm  
Generalized (6x6) diffusion matrix: (Dtt Dtr)  
(Drt Drr)  
4.335E-06 -7.138E-09 -1.327E-07 1.299E-01 -1.848E-01 -2.874E-01  
-7.023E-09 4.538E-06 1.781E-07 -1.845E-01 -4.769E-01 -1.174E-01  
-1.328E-07 1.782E-07 4.754E-06 -2.874E-01 -1.177E-01 4.969E-01  
1.299E-01 -1.845E-01 -2.874E-01 3.232E+08 -2.214E+07 -3.601E+07  
-1.848E-01 -4.769E-01 -1.177E-01 -2.214E+07 3.740E+08 8.775E+07  
-2.874E-01 -1.174E-01 4.969E-01 -3.601E+07 8.775E+07 4.956E+08

ROTATIONAL DIFFUSION TENSOR

3.232E+08 -2.214E+07 -3.601E+07  
-2.214E+07 3.740E+08 8.775E+07  
-3.601E+07 8.775E+07 4.956E+08

Anisotropic rotational diffusion

Eigenvalue (s<sup>-1</sup>) -- Eigenvector -----  
1 3.147E+08 0.9612 0.2690 0.0608  
2 3.286E+08 -0.2062 0.8472 -0.4896  
3 5.494E+08 -0.1832 0.4581 0.8698

Anisotropic rotational diffusion

Dx, Dy, Dz (s<sup>-1</sup>) -- Eigenvector -----  
Dz 5.494E+08 -0.1832 0.4581 0.8698  
Dy 3.147E+08 0.9612 0.2690 0.0608  
Dx 3.286E+08 -0.2062 0.8472 -0.4896

Rotational diffusion coefficient: 3.976E+08 s<sup>-1</sup>

Rotational diffusion anisotropy: 2.281E+08 s<sup>-1</sup>

Relaxation time (1): 5.183E-10 s

Relaxation time (2): 4.680E-10 s

Relaxation time (3): 4.590E-10 s

Relaxation time (4): 3.520E-10 s

Relaxation time (5): 3.519E-10 s

Harm. mean relax. (correlation) time: 4.192E-10 s

Iflag= 1 N-H bonds, implicit hydrogens (not given in pdb)

----- Calculation for 11.74 Tesla -----

NMR DATA AND CONSTANTS

Magnetic field: 11.74 Tesla

Gyromagnetic ratio of 1H : 2.675E+04 rad.s<sup>-1</sup>.G<sup>-1</sup>

Gyromagnetic ratio of X : -2.713E+03 rad.s<sup>-1</sup>.G<sup>-1</sup>

Angular resonance frequency of 1H : 3.140E+09 rad.s<sup>-1</sup>

Resonance frequency of 1H : 500. MHz

Angular resonance frequency of X : 3.185E+08 rad.s<sup>-1</sup>

Internuclear H-X distance : 1.020E-08 cm

Chemical shielding anisotropy : -1.600E-04

Residue-specific T1, T2, RNOE, etc, have been calculated for 0 residues (out of 1), Mean <T1/T2>= NaN NO. IND-CHAIN AMIN T1/T2 NABLA

T1 T2 RNOE TMEAN TINI 1 1 UNK

TABLE OF SPECTRAL DENSITIES

|           | 0         | WX        | WH-WX     | WH        | WH+WX     | 0.86WH    |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| NO.       | 0.000E+00 | 3.185E+08 | 2.822E+09 | 3.140E+09 | 3.459E+09 | 2.732E+09 |
| IND-CHAIN | AMIN      | J(0)      | J(WX)     | J(WH-WX)  | J(WH)     |           |
| J(WH+WX)  | J(0.86WH) | 1         | 1         | UNK       |           |           |