Supplemental data to the manuscript:

Calcium-induced folding of intrinsically disordered Repeat-in-Toxin (RTX) motifs via changes of protein charges and oligomerization states

Ana Cristina Sotomayor Pérez, Daniel Ladant and Alexandre Chenal

Supplemental Materials and Methods

Polypeptides production and purification

The RTX polypeptides production and purification have been described elsewhere (1). Briefly, the RTX polypeptides were overproduced in *E. coli* and purified by chromatography on a nickel HiTrap chelating column (elution in 500 mM imidazole, 20 mM Hepes, 200 mM NaCl, pH 7.4), then by a size exclusion chromatography on Sephacryl S200 (20 mM Hepes, pH 7.4, 100 mM NaCl) and finally by ion exchange chromatography on Q sepharose (elution with 20 mM Hepes, 500 mM NaCl pH 7.4) at room temperature (1). Ethylenediaminetetraacetic acid (EDTA) at a final concentration of 2 mM was added to samples before desalting on prepacked G25SF against 20 mM Hepes, 100 mM NaCl pH 7.4 (for storage at –20 °C) or against 20 mM NH_4HCO_3 and lyophilized. The different RTX polypeptides were more than 95% pure as judged by SDS-PAGE. The integrity and identity of the samples was confirmed by N-terminal sequencing and by mass spectrometry (SELDI-TOF-MS model PCS 4000, Ciphergen). Polypeptide concentrations were determined by UV spectrophotometry using molar extinction coefficients (at 280 nm) of 18 000 $M^{-1}cm^{-1}$ for NRC_L and RC_L, 16 500 $M^{-1}cm^{-1}$ for NRC_S and RC_S, 9 600 $M^{-1}cm^{-1}$ for NR and 9 500 $M^{-1}cm^{-1}$ for R.

Quasi elastic light scattering

QELS experiments were done with two different instruments: (i) a DynaPro MS800 (Wyatt) that monitored the scattered light at 90° using the Protein Solution Dynamics software version 6.2.05 and (ii) a NanoZS (Malvern) which monitors the backscattered light at 173° using the dedicated software. Buffers and samples were filtrated on 0.2 µm filters prior to acquisition and RTX polypeptide concentration was 55 µM. A microcuvette of dimensions 3*8.5 mm (105.251-QS) was loaded with 150 µL of RTX polypeptides in buffer A. Samples were equilibrated for 10 minutes in the cell compartment before acquisition. Acquisition time was 10 sec, with an interval time of 1 sec. At least 30 acquisitions were averaged to produce a data

collection. A set of three independent data collections was obtained for each experimental condition. Processing of the data acquired with the DynaPro MS800 instrument has been described elsewhere (1). Processing of the data acquired on the NanoZS was done using the Mean narrow band and Protein Analysis procedures. Similar results were obtained with both instruments.

Circular dichroism spectroscopy

CD spectra were recorded on an Aviv circular dichroism spectrometer model 215 equipped with a water-cooled Peltier unit. CD experiments were carried out at 25 °C and at a scan rate of 0.5 nm/sec (step: 0.5 nm; integration time: 1 sec) with a time constant of 100 msec and a bandwidth of 1 nm. Rectangular quartz Suprasil cells of 0.1 and 10 mm path-lengths (106-QS and 114B-QS, Hellma) were used for recording CD signals in far-UV and near-UV regions, respectively. Each far-UV and near-UV CD spectrum represents the average of at least 5 scans. Buffer A was used as blank in far-UV and near-UV regions and its spectrum was subtracted to protein CD spectra. The CD units used are the mean residue ellipticity for CD spectra in the far-UV range $[\theta]_F$ ((deg x cm^2)/(dmol x res)) and the molar ellipticity for CD spectra in the near-UV range $[\theta]_N$ ((deg x cm^2)/dmol), respectively.

Size exclusion chromatography followed by intrinsic viscosity and molecular mass measurements

Size exclusion chromatography (SEC) experiments were performed using a Superdex 200 column (GE Healthcare). The system was controlled by a GPCmax module and connected on-line to a Tetra Detector Array (TDA) model 302 (Viscotek Ltd., a Malvern Company). The oven of the TDA contains (i) a static light scattering cell with two photodiode detectors at 7° for low angle (LALS) and at 90° for right angle laser light scattering (RALS), (ii) a deflection refractometer, (iii) a photometer and (iv) a differential viscometer.

The general procedures defined by Viscotek were followed as previously described elsewhere (2) (3) (4). All solutions were filtered on 0.2 µm filters and allowed to equilibrate at 22 °C prior to running. The SEC experiments were performed at 22 °C while the detections in the TDA oven were done at 25 °C. All experimental sequences contained injections of NIST-1923 PolyEthylene Oxide 22-KDa (PEO from Viscotek PolyCal^{TM} TDS-PEO-N, 100 µL at 4 g/L; intrinsic viscosity: 39 mL/g in our experimental conditions) and BSA (SIGMA A0281, 200 µL at 2.4 mg/mL; intrinsic viscosity: 4 mL/g) used for TDA internal constant

calibrations, followed by the samples of apo-RC_S or holo-RC_S (at least three injections of different concentrations). The buffer used was buffer A for apo-RC_S and buffer A supplemented with 5 mM $CaCl_2$ for holo-RC_S. Polypeptide concentrations were determined using both the photometer and the deflection refractometer. RALS and LALS data combined with the protein concentrations provided the molecular mass, while the differential viscometer measurements, in combination with the protein concentrations provided the intrinsic viscosity. All data were acquired and processed using the Viscotek's Omnisec software. Samples of apo-state of NRC_L, RC_L (Fig. 2) and R polypeptides were similarly analyzed by SEC-TDA. The data indicate that all apo-states were monomeric and exhibits a high intrinsic viscosity: $[\eta]$ is 15 ± 0.5 mL/g for NRC_L, 13.7 ± 0.2 mL/g for RC_L and 12.6 ± 0.2 mL/g for R.

Analytical ultracentrifugation

Velocity experiments were performed at 25 °C on a Beckman XL-I analytical ultracentrifuge (Beckman Coulter) in an AN60-Ti rotor. Detection of the polypeptide concentration as a function of radial position and time was performed by optical density measurements at a wavelength of 276 nm. RC_S polypeptide samples (400 µL at 20 µM) were loaded in a 1.2 mm-thick two channels aluminium centerpiece and spun at 60000 rpm. Data were analyzed with the Sedfit 11.3 software (5) using a continuous size distribution c(s) model to extract the sedimentation coefficient, S. We used the Svedberg relation

$S = \frac{M(1-\bar{\nu}\rho)}{6\pi\eta_w R_H N_A} = \frac{D_t M(1-\bar{\nu}\rho)}{RT}$ to calculate the translational diffusion coefficient D_t and the hydrodynamic radius R_H using the experimental values of the molecular mass M and the sedimentation coefficient S.

Protein shape and hydration

For the determination of protein shape and hydration, we used a previously described procedure (1) (2,3) (4). Briefly, the protein shape and hydration can be experimentally determined from the molecular mass (M), the intrinsic viscosity $[\eta]$, the hydrodynamic radius R_H and the partial specific volume $\bar{\nu}$ of the protein under study (see below for details). The viscosity increment ν (also called the Simha-Einstein hydrodynamic function) is related to the shape of the protein and can be calculated from the Einstein's viscosity relation: $M[\eta] = \nu V_H N_A$, where V_H is the hydrodynamic volume defined by $V_H = 4\pi R_H^3/3$, and N_A the Avogadro number. The viscosity increment ν provides an estimation of the ratio of the

lengths of the semi-axis a and b of an ellipsoid of revolution describing the envelope of the investigated protein (6).

The intrinsic viscosity in a defined solvent mainly depends on the shape and on the hydration of the protein: $[\eta] = \nu(\bar{\nu} + \delta/\rho)$, from which the hydration parameter is calculated: $\delta = (([\eta]/\nu) - \bar{\nu})\rho$. The hydration is also calculated from the relation $M(\bar{\nu} + \delta/\rho) = V_H N_A$, which is independent from the intrinsic viscosity. This relation is obtained by substituting the intrinsic viscosity by its expression $[\eta] = \nu(\bar{\nu} + \delta/\rho)$ into the Einstein's viscosity relation. The time-averaged apparent hydration parameter of the protein, δ (gram of water per gram of protein), includes (*i*) the water molecules bound to the protein and (*ii*) the water molecules dragged by the diffusion of the protein.

We also used the Perrin equation,

$P = (f/f_0)\Big/\left[\left((\delta/\bar{\nu}\rho) + 1\right)^{1/3}\right]$, to calculate the Perrin hydrodynamic function P, which provides an alternative estimation of the semi-axial ratio a/b of the protein envelope (6-8). We used the hydration estimated from the Einstein viscosity relation and the frictional ratio (determined by QELS or AUC) to obtain the P shape factor. The partial specific volume $\bar{\nu}$ was determined by SEDNTERP software.

Steady-state fluorescence anisotropy spectroscopy.

Measurements were performed using a FP-6200 spectrofluorimeter (Jasco, Japan) equipped with a Peltier-thermostated cell holder, using 2 mL samples in 1 cm path-length quartz cells (101-QS from Hellma). A bandwidth of 5 and 10 nm was used for the excitation and emission beams respectively. RTX polypeptides concentrations were 2 µM. The emission wavelength was fixed at 350 nm. The excitation spectra were recorded at 25 °C from 270 to 290 nm at a scan rate of 125 nm.min^{-1}. Experiments were performed with vertically (V) and horizontally (H) polarized beam light using FDP-223 polarizers at both excitation (x) and emission (m) apertures. Anisotropy (r) values were calculated as follows:

$$r = \frac{V_x V_m - \left(\frac{H_x V_m}{H_x H_m}\right) V_x H_m}{V_x V_m + 2\left(\frac{H_x V_m}{H_x H_m}\right) V_x H_m}$$

Electrophoretic mobility

Electrophoretic mobility experiments were performed at 25 °C on a NanoZS instrument (Malvern) which monitors the forward light scattering at an angle of 17°. Buffers and samples were filtrated on 0.2 µm filters prior to acquisition. RTX polypeptide concentrations ranged between 70 to 130 µM and were diluted in 20 mM Hepes, 20 mM NaCl, pH 7.4. A dedicated cuvette DTS1060 and a microcuvette ZEN1010 were used for the acquisition of the electrophoretic mobility values using the fast field reversal mode. The quality of the instrumentation was checked using the standard DTS1230. For protein sample measurements, the intensity was always kept below 5 mA irrespectively of the applied voltages (from 30 to 100 V) as the conductivities of the various protein samples were all below 0.5 $mS.cm^{-1}$, due to the low ionic strength of the buffers used in this study. Before starting electrophoretic mobility experiments, a series of QELS experiments was run to determine the R_H of the species and check their homogeneity (with at least, more than 90% of the scattered intensity corresponding to the specie of interest). For each sample, five to ten independent measurements (one run each) were acquired in fast field reversal mode. The quality criteria to keep a measurement were based on: (i) the zeta quality factor (signal-to-noise ratio of the frequency shift); (ii) the mean count rate that should not change throughout the duration of the data acquisition; (iii) the quality of the phase plot (radian amplitude and frequency) and the Fourier-transformed of the phase plot. The acceptable electrophoretic mobility measurements were averaged and the standard deviation was computed. These data were used to generate the electrophoretic mobility distribution of each state (apo and holo) of both proteins, RC_L and RC_S.

In an attempt to estimate the number of charges of the studied RTX-proteins, we used the following empirical relation described by Basak and Ladisch (9):

$$N_C = \left(\left(\left(f/f_0\right) \times \mu_e\right) + B\right) M^{2/3} A^{-1},$$

where Nc is the number of charges (the protein valence), f/f_0 the frictional ratio, μ_e the electrophoretic mobility, M is the protein molecular mass, A and B are the values of the slope and the y-intercept, respectively. This relation, which has been obtained with a large data set of proteins (see (9) for details), indicates that the electrophoretic mobility (the velocity of a protein in an given electric field) is proportional to the number of charges (Nc) and inversely correlated to the frictional ratio (f/f_0) between the solvent and the protein, and to the 2/3 power of molecular mass (M).

Supplemental References

1. Sotomayor Perez, A. C., Karst, J. C., Davi, M., Guijarro, J. I., Ladant, D., and Chenal, A. (2010) *J Mol Biol* **397**, 534-549
2. Chenal, A., Guijarro, J. I., Raynal, B., Delepierre, M., and Ladant, D. (2009) *J Biol Chem* **284**, 1781-1789
3. Bourdeau, R. W., Malito, E., Chenal, A., Bishop, B. L., Musch, M. W., Villereal, M. L., Chang, E. B., Mosser, E. M., Rest, R. F., and Tang, W. J. (2009) *J Biol Chem* **284**, 14645-14656
4. Karst, J. C., Sotomayor Perez, A. C., Guijarro, J. I., Raynal, B., Chenal, A., and Ladant, D. (2010) *Biochemistry* **49**, 318-328
5. Schuck, P. (2000) *Biophys J* **78**, 1606-1619
6. Harding, S. E., and Colfen, H. (1995) *Anal Biochem* **228**, 131-142
7. Perrin, F. (1936) *J. Phys. Rad.* **7**, 1-11
8. Squire, P. G., and Himmel, M. E. (1979) *Arch. Biochem. Biophys.* **196**, 165-177
9. Basak, S. K., and Ladisch, M. R. (1995) *Anal Biochem* **226**, 51-58

Supplemental figures

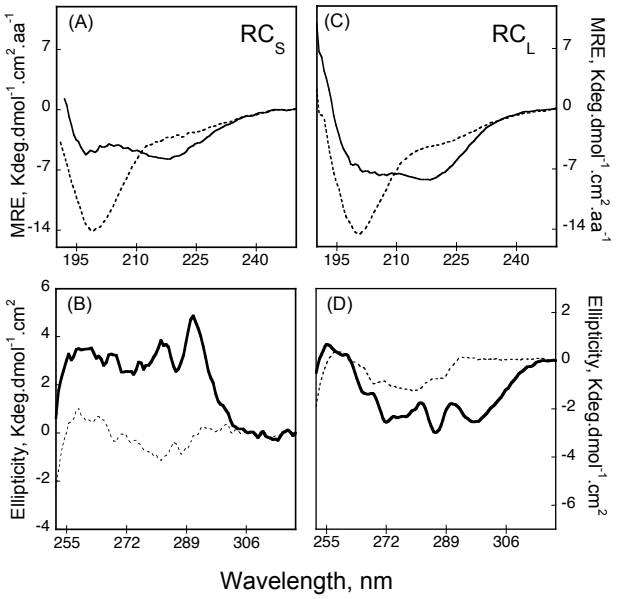

Figure S1

Supplemental figure legends

Figure S1. CD spectroscopy of RC_S and RC_L polypeptides. Far-UV CD spectra of RC_S (**A**) and RC_L (**C**) polypeptides in the absence (dotted line) or in the presence of calcium (2 mM for RC_L and 5 mM for RC_S, continuous line). Near-UV CD spectra of RC_S (**B**) and RC_L (**D**) in the absence (dotted line) or in the presence of calcium (2 mM for RC_L and 5 mM for RC_S, continuous line). Polypeptide concentrations are 30 μM for far-UV CD and 60 μM for near-UV CD. Experimental conditions: buffer A, 25 °C.

Supplemental Table

Table S1. Hydrodynamic parameters of RC_S polypeptide in the absence or presence of calcium.

Parameters	RC_S	RC_S + Calcium
Retention volume [a] mL	18.5	18.4
Intrinsic viscosity [a] (mL/g)	14.7 ± 0.2	6.4 ± 0.1
Molecular mass [a] (kDa)	15.1 ± 1.1	42.5 ± 1.2
D_t [b] (x 10^{-7})	8.8 ± 0.2	7.6 ± 0.1
Sedimentation coefficient [b] (S)	1.46 ± 0.1	3.65 ± 0.1
R_H [b] (nm)	2.9 ± 0.1	3.1 ± 0.2
R_0 anhydrous (nm)	1.58	2.30
Frictional ratio [b] (f/fo)	1.8 ± 0.1	1.35 ± 0.1
Hydration [c] (g/g)	3.0 ± 0.2	1.1 ± 0.15
Viscosity increment [c]	3.9 ± 0.3	3.4 ± 0.2
Perrin value [c]	1.020	1.030
Partial specific volume [c] (mL/g)	0.7115	0.7115

[a] Experimental data from SEC-TDA
[b] Experimental data from AUC
[c] Hydrodynamic parameters calculated as described in Material and Methods.