Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics
This journal is © The Owner Societies 2013

Electronic Supplementary Information

Initiation of Assembly of Tau(273-284) and its ΔK280 Mutant: An Experimental and Computational Study

Luca Larini[a,b,‡], Megan Murray Gessel[b,‡,‖], Nichole E. LaPointe[c,‡], Thanh D. Do[b,‡], Michael T. Bowers[b], Stuart C. Feinstein[c], Joan-Emma Shea[a,b,*]

[a] *Department of Physics, University of California at Santa Barbara, Santa Barbara, California, 93106, U.S.A. E-mail shea@chem.ucsb.edu*
[b] *Department of Chemistry and Biochemistry, University of California at Santa Barbara, Santa Barbara, California, 93106, U.S.A.*
[c] *Neuroscience Research Institute and Department of Molecular, Cellular and Developmental Biology, University of California at Santa Barbara, California, 93106, U.S.A.*
[‡] *These authors contributed equally to the manuscript.*
[*] *Corresponding author.*
[‖] *Current address Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee*

		IM-MS		REMD	
		R2/wt	R2/ΔK280	R2/wt	R2/ΔK280
n	z	ccs	Ccs	ccs	ccs
1	2	310	288	312-344, (317.4)	298-313, (291)
1	3	336	n/a		
2	3	516, 528, 545	482	Compact:501-571, Extended:582, (575,556)	Parallel:525,490, Antiparallel:531, Compact: 484, (455,473,496)
2	4	575, 600	n/a		
3	4	687, 715	636, 667, 698	n/a	n/a
4	6	964	n/a	n/a	n/a

Table S1 Comparison between experimental and theoretical collisional cross sections (CCSs). For experimental cross sections, in cases where more than one conformation of a given oligomer (n) was present in the experiment, two cross sections are reported. Theoretical CCSs in vacuum are reported in parentheses. All units are $Å^2$.

This journal is © The Owner Societies 2013

Computational details

Simulation details

Simulations were performed using the force field and parameters as described in the manuscript.

In order to find the equilibrium volume used for the REMD simulations, a prior equilibration in the NPT ensemble was performed, at constant temperature T=300K and pressure P=1bar, using the Berendsen thermostat (time constat τ_T=0.1ps) and barostat (time constant τ_P=1ps). The number of water molecules was the same for all the systems studied and equal to n_w=7084. The total charge of each system under examination was neutralized using Cl^- or Na^+ ions. The systems were equilibrated in NPT ensemble for at least 2ns. The final equilibrium volume differed slightly among different cases, with the box edges being (in nm): 6.008 (monomer R2/wt, R2/ASPH and R2/LYS), 6.027 (dimer R2/wt), 6 (monomer R2/ΔK280), 6.04(dimer R2/ΔK280).

Once the equilibrium volume was reached, the systems were further equilibrated in the NVT ensemble using the Nosé-Hoover thermostat (as described in the manuscript) for at least 10ns.

For the REMD simulations, different sets of temperature were tested to reach the desired exchange ratio of ~25%. The initial set of temperature was taken from the website http://folding.bmc.uu.se/remd/index.php, according to the work by Patriksson and van der Spoel[1], and then manually fine-tuned. The final set of temperature, in K, for each case is :

- Monomer (R2/wt and R2/ΔK280): 290, 292.5, 295, 297.5, 300, 302.5, 305, 307.5, 310, 312.5, 315, 318, 321, 324, 327, 330, 333, 336, 339, 342, 345, 348, 351, 354, 357, 360, 363, 366, 369, 372, 375, 378, 381, 384.5, 388, 391.5, 395, 398.5, 402, 405.5, 409, 412.5, 416, 420, 424, 428, 432, 436, 440, 444, 448, 452, 456, 460, 464, 468, 472, 476, 480, 484, 488, 492
- Monomer (R2/LYS, R2/ASPH): 290, 292.5, 295, 297.5, 300, 302.5, 305, 307.5, 310, 312.5, 315, 317.5, 320.5, 323.5, 326.5, 329.5, 332.5, 335.5, 338.5, 341.5, 344.5, 347.5, 350.5, 353.5, 356.5, 359.5, 362.5, 365.5, 368.5, 371.5, 374.5, 377.5, 381, 384.5, 388, 391.5, 395, 398.5, 402, 405.5, 409, 412.5, 416.2, 420.2, 424.2, 428.2, 432.2, 436.2, 440.2, 444.2, 448.2, 452.7, 457.2, 461.7, 466.4, 471.1, 475.8, 480.5, 485.2, 489.9, 494.6, 500.8
- Dimer (R2/wt): 290.8, 293.1, 295.4, 297.7, 300, 302.5, 305, 307.5, 310, 312.5, 314.5, 317, 319.5, 322, 324.7, 327.4, 330.1, 332.8, 335.5, 338.2, 341.2, 344.2, 347.2, 350.2, 353.2 , 356.2
- Dimer (R2/ΔK280): 290.8, 293.1, 295.4, 297.7, 300, 302.5, 305, 307.5, 310, 312.5, 315.1, 317.6, 320.1, 322.6, 325.3, 328, 330.7, 333.4, 336.1, 338.8, 341.8, 344.8, 347.8, 350.8

Another set of simulations was performed in vacuum to mimic the conditions that the molecules experience during the mass spect rometry. In this way we can make sure to identify structures that appear in vacuum, but that are otherwise not present in solution. These simulations follow the protocol described in the manuscript. All the simulations in vacuum are 1μs long, and only the final 400 ns are analyzed. The initial structures are randomly singled out from the simulations in solution. The acceptance ratio for the REMD simulation is ~25% and the set of temperature was selected as explained above for simulations in presence of solvent. Restraints were imposed to avoid isomerization of the peptide bond at the highest temperatures. The final set of temperatures used is (in K):

- Monomer (R2/wt and R2/ΔK280): 270.00, 300, 333, 370.46, 411.05, 455, 502.64, 554.31, 610.36, 671.15, 742.12, 818.65, 901.29, 993.56, 1095.04, 1214, 1350, 1505, 1670, 1860
- dimer (R2/wt): 279.00, 300, 323, 345.46, 369.05, 394.5, 424.1, 456.31, 487.36, 519.15.
- dimer (R2/ΔK280): 279.00, 300, 323, 344.46, 368.05, 393.5, 423.1, 455.31, 487.36, 521.15.

Computing the collisional cross-sections

Collisional cross-sections (CCS) were computed for both the gas phase conformations of the peptide as well as for the conformations in solution. In the latter case, the conformations were dehydrated to better match the experimental condition of the mass spectrometry. The dehydration consisted in minimizing the energy of the peptide chain in vacuum using a steepest descent algorithm.

Relevant conformations were singled out using the Daura algorithm[2] described in the manuscript. This algorithm allows us to identify clusters of structures with similar conformations and their abundance. From each cluster, a set of random conformations

[1] A. Patriksson and D. van der Spoel, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2073

[2] X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, and A.E. Mark, *Angew. Chem. Int.*, 1999, **38**, 236.

This journal is © The Owner Societies 2013

is selected and the CCSs computed. The average values are reported in Table S1.

Despite working very well in the gas phase, this strategy is not optimal for solution conformations. In fact, in the latter case the peptide populates multiple conformational with a very low probability. As a consequence, the Daura algorithm generates a myriad of small clusters. In such situation a different approach was adopted.

In the case of solution phase simulation, however, the peptide populates multiple conformations with a low probability, so that the identification of predominant clusters is not possible. For this reason a different approach was adopted. The structures were classified according to their geometric properties, namely the end-to-end distance (which is a measure of how elongated the backbone is) and the radius of gyration (which measures how compact a conformation is, not just the backbone but the side chains as well). In this way, conformations with an end-to-end distance that differs more than 0.25nm are associated to different cluster. Afterwards, each of these clusters is further split into smaller clusters if its conformations differ more than 0.05nm in the radius of gyration. This procedure is equivalent to drawing a grid in the R_{ee} vs R_g plots shown in the manuscript and picking one conformation for each point of the grid.

The procedure just outlined shows that the CCS are proportional to increasing R_{ee} and increasing Rg, (Fig. S1,Fig. S2) proving that these quantities properly describe the conformations adopted by the peptide. However, as explained above, to better mimic the experimental conditions, these structures are minimized using a steepest descent algorithm in the gas phase. Following this procedure, we have found that if the conformation is compact in solution, then its cross-section will not change much. However, for very extended structures (high value of R_{ee} and R_g), this is not the case any longer.

As a final remark, it should be noted that the above procedure is useful to define a range of possible CCS, not an average value. Even though an average value could be computed, such average value cannot be associated to any predominant structures. For this reason we have decided to report the range, which better represents a heterogeneous set of conformations.

On the other hand, both monomers at varying protonation and dimers in solution show some predominant clusters, so that the CCS computed from such clusters are reported. A peculiar situation arises in the wild type dimer in solution, where a dominant cluster is observed, even though the majority of the conformations adopt disordered structures. In this case a hybrid approach was adopted, where the CCSs were firstly computed for the predominant clusters. At this point the conformations belonging to this cluster were removed from the pool of all the conformations. The CCSs were computed for these remaining conformations using their geometric properties (Fig. S3).

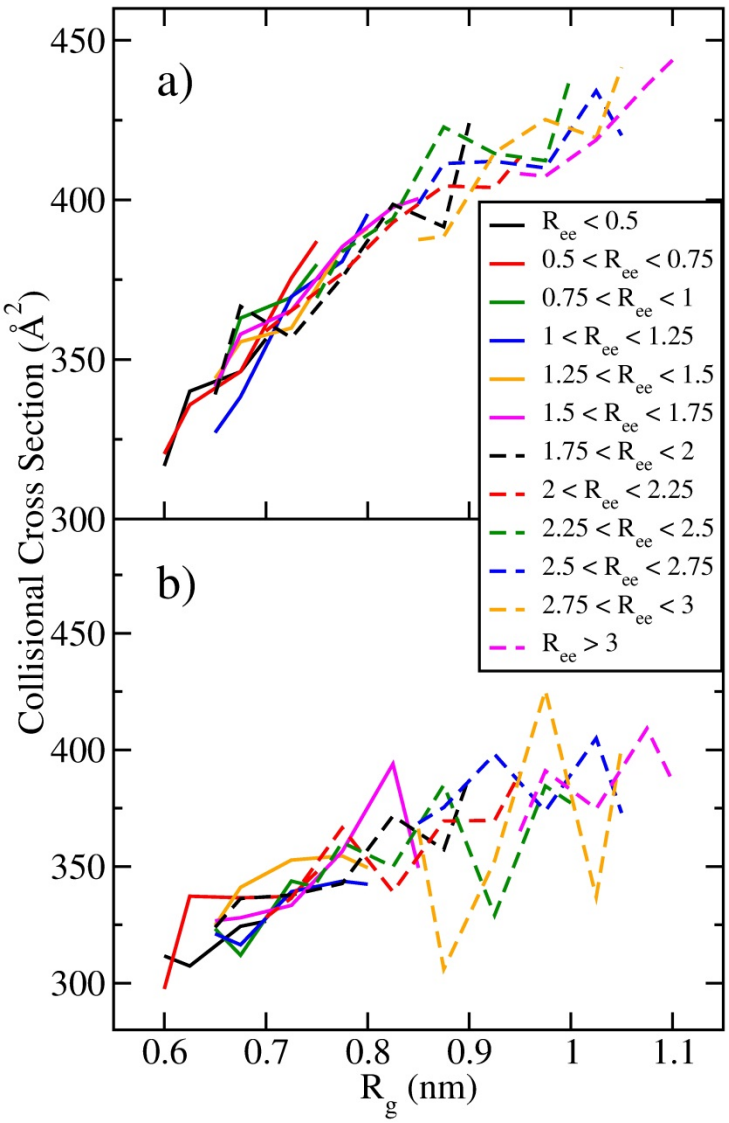


Figure S1 Collisional cross section (CCS) computed using the REMD trajectories as explained in the "Methods" section. The structures are classified according to their end-to-end distance (R_{ee} in nm) and their radius of gyration (R_g). The plot shows the CCS in solution of the wild-type monomer. a) CCS computed using the structures from the solution before minimization. b) The same CCS computed after minimization in the gas phase. Minimization in the gas phase is required to better mimic the experimental conditions in which the IM-MS experiments are performed. If we consider only the region with highest probability (Fig. 6 (a), $0.5 < R_{ee} < 1.5$, $0.61 < R_g < 0.7$), then the most likely cross-sections are in the range 312 – 344. For simulations performed in gas phase, the average CCS is 317.4, with a standard deviation of 3.9.

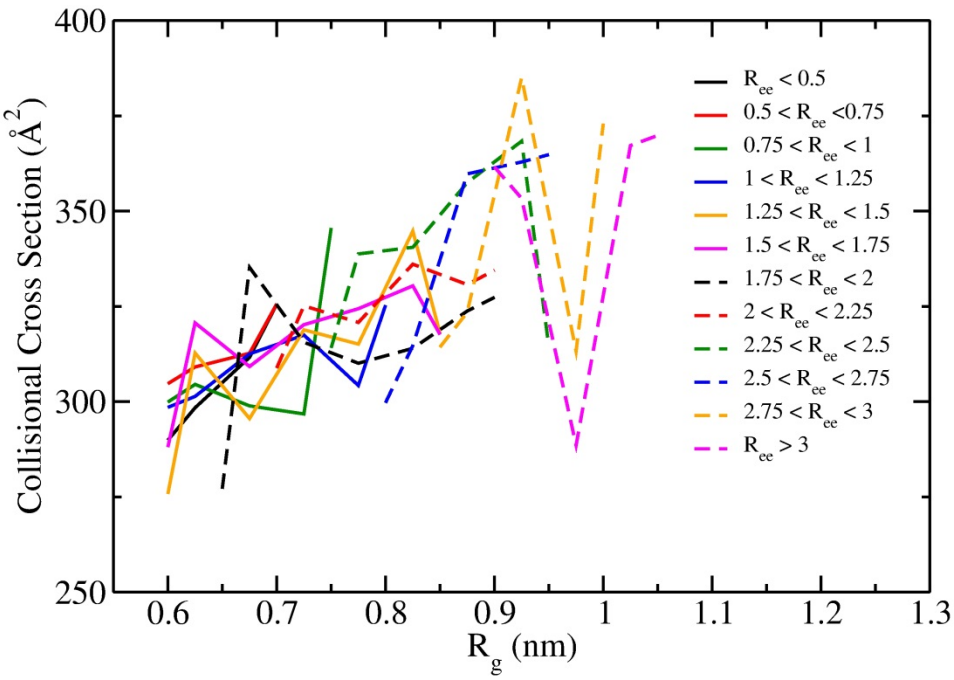
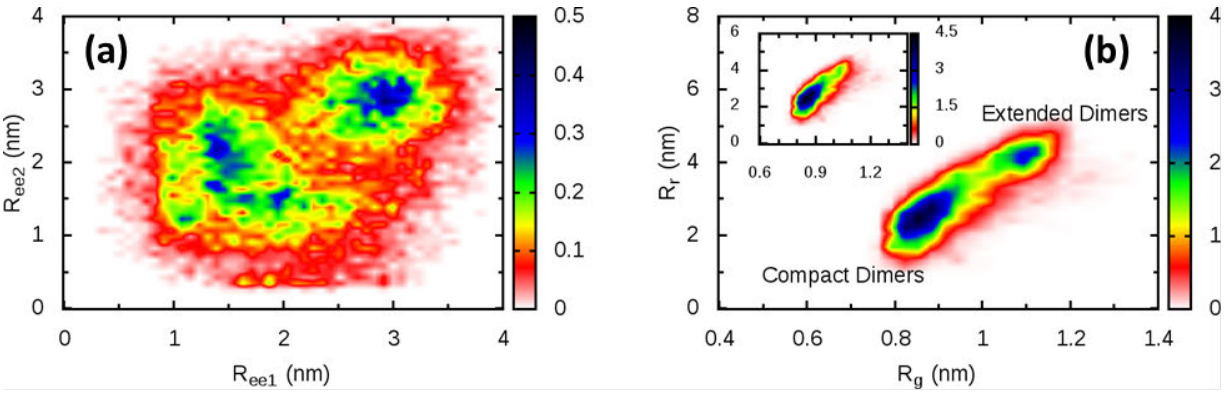


Figure S2 The monomer of the mutant R2/ΔK280 peptide populates multiple conformations, without any of them being the dominant one. For this reason the ensemble of conformation was split according to their geometric properties, namely end-to-end distance R_{ee} (in nm) and radius of gyration R_g. For each subgroup constructed in this way a random conformation was selected and the corresponding cross-section computed. A comparison between this Figure and Fig. 6(b), shows that the most likely cross-section is in the range 298-213 (corresponding to 0.5 < Ree < 1.25; 0.58 < Rg < 0.66, i.e. the “blue spot in Fig. 6(b)).



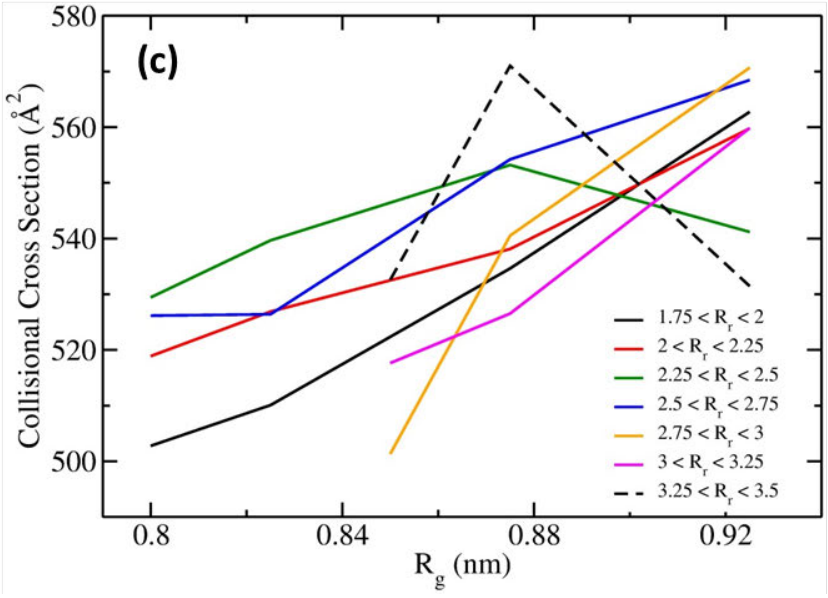


Figure S3 CCSs for the wild-type dimer is solution. a) The distribution of end-to-end distance of the first chain (R_{ee1}) versus the other (R_{ee2}) shows two fuzzy regions roughly located on the diagonal of the plot. This suggests that the two chains tend to extended together. As a consequence a generalized coordinate that lies on the diagonal of this plot is appropriate to describe the extension of both chains at the same time. b) For these reasons we have plotted the generalized coordinate $R_r = \sqrt{R_{ee1}^2 + R_{ee2}^2}$ versus the radius of gyration (R_g) of the whole aggregate. The generalized coordinate is equivalent to fix the radius of a circle with center in the origin of a) and count the number of conformations that lay within a small distance to it. In this way it is possible to clearly distinguish two populations within the dimer. The inset shows the same plot after removing the conformations belonging to cluster A in Fig. 8 of the manuscript. c) The CCSs of the compact dimers of panel b) at $1.75 < R_r < 3.5$and $0.8 < R_g < 0.95$ ("blue spot" in b)).

This journal is © The Owner Societies 2013

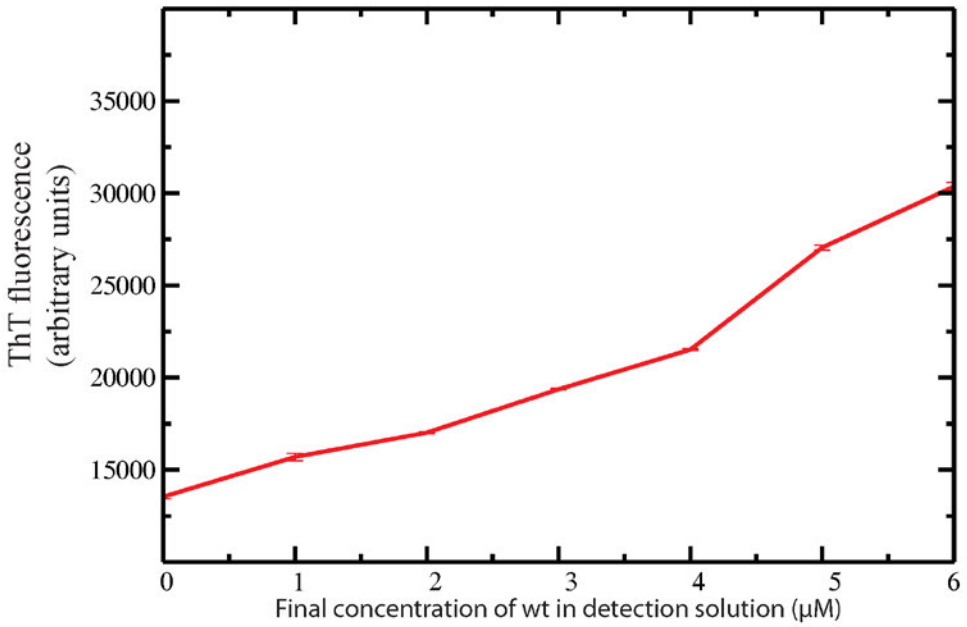


Figure S4 To determine the linear detection range of the ThT assay, we used R2/wt (50 μM) that had aggregated > 24 h in ammonium acetate buffer in the presence of heparin. We added this fully aggregated sample to ThT detection solution at final peptide concentrations ranging from 1-6 μM, and measured the resulting ThT fluorescence signal.

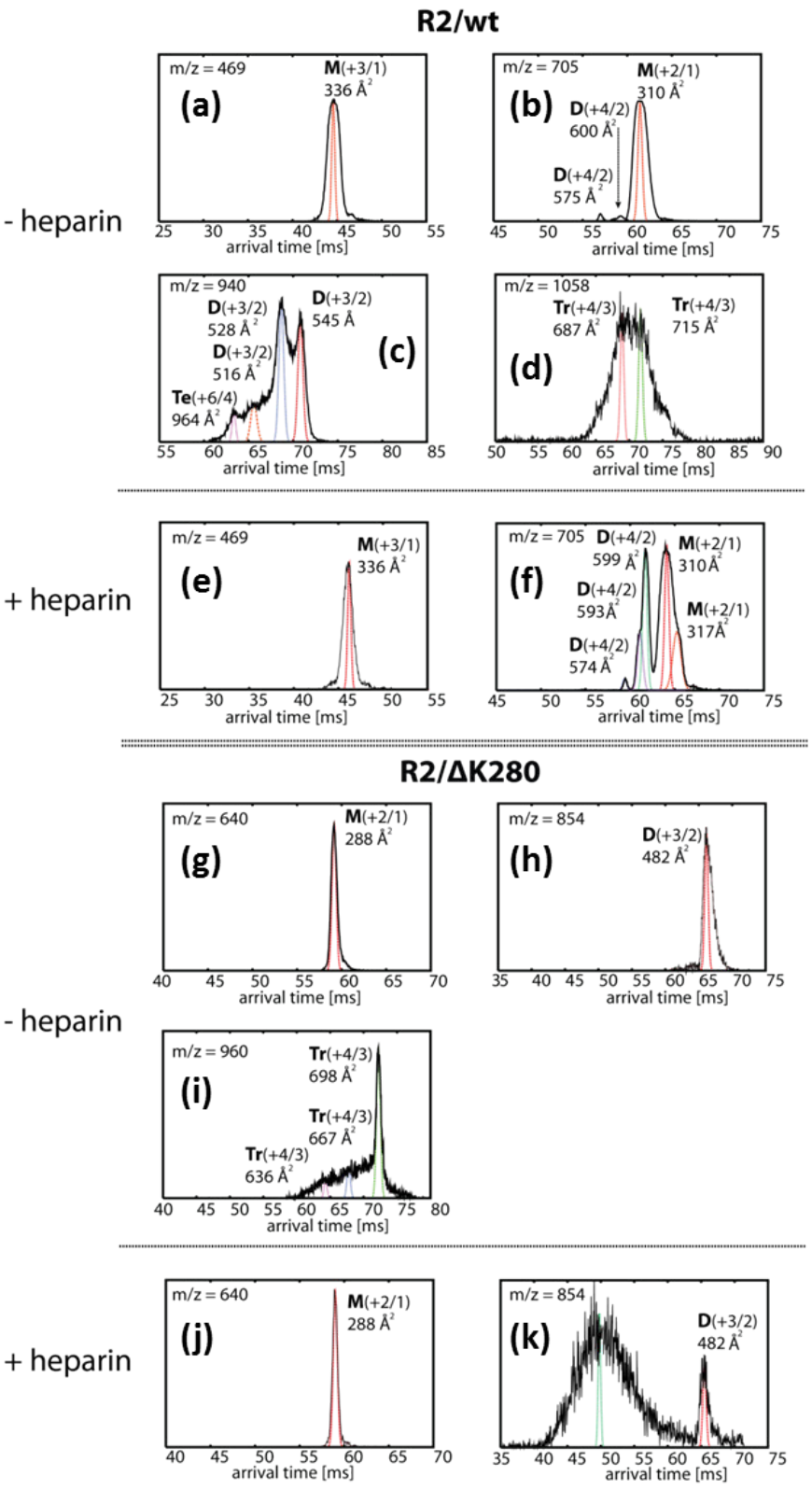


Figure S5 Representative ATDs of R2/wt and R2/ΔK280 in 20 mM ammonium acetate buffer.

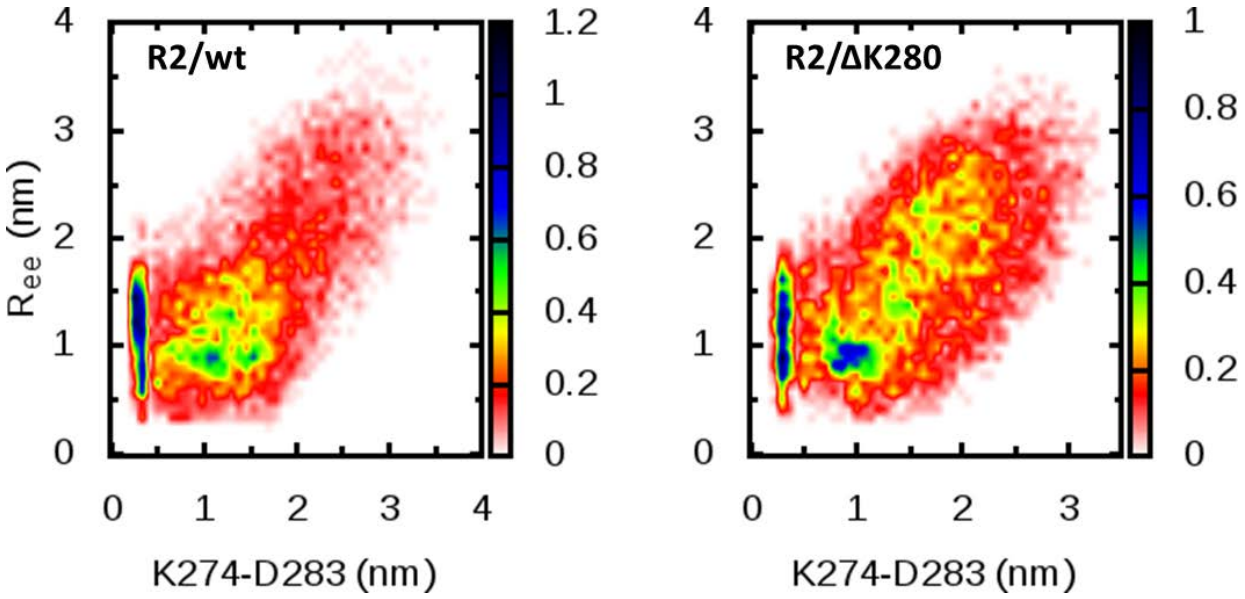


Figure S6 Probability of finding a conformation according to its end-to-end distance and the distance between K274 and D283 for both R2/wt and mutant R2/ΔK280.

This journal is © The Owner Societies 2013

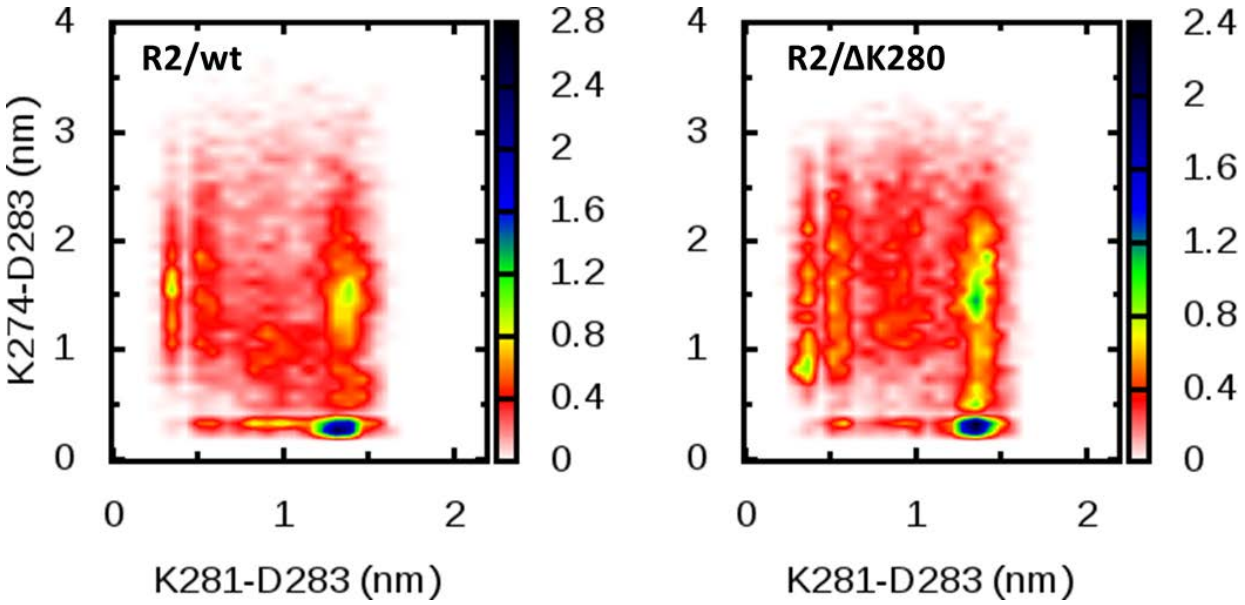


Figure S7 The probability of finding a salt bridge between pairs of residues for both wild-type (R2/wt) and mutant (R2/ΔK280). The presence of salt-bridges can be inferred by the close proximity of residues.

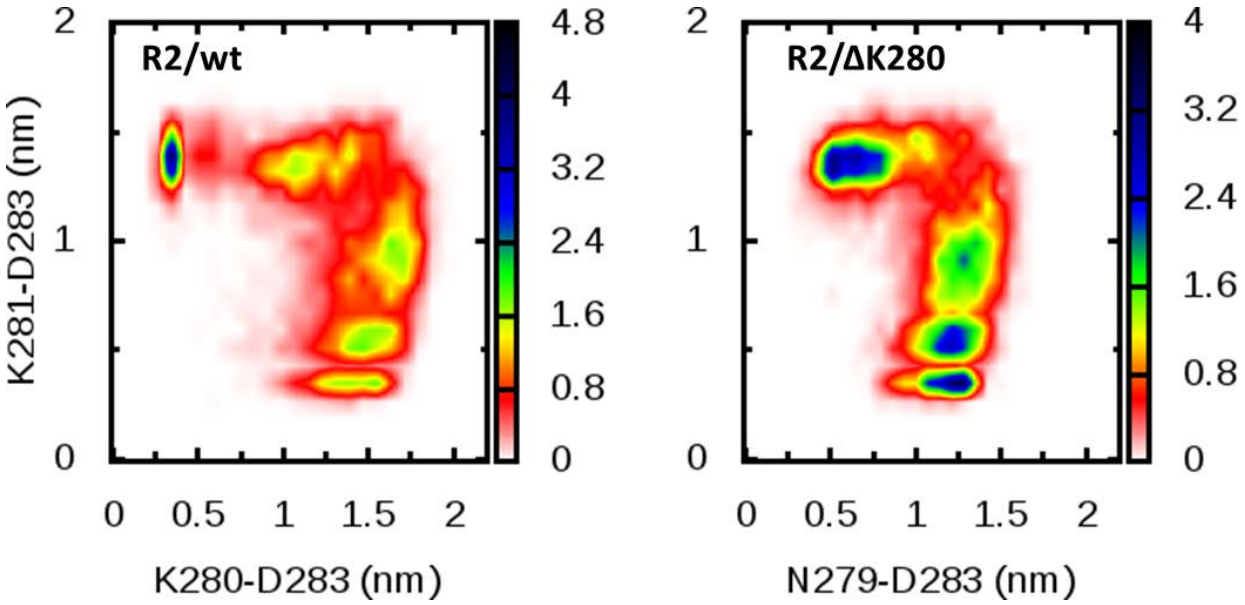


Figure S8 This plot shows the probability of having pairs of residues in close proximity, which indicates the presence of salt bridges (K281-D283 and, for R2/wt only, K280-D283) or hydrogen bonds between side chains (N279-D283, mutant only).

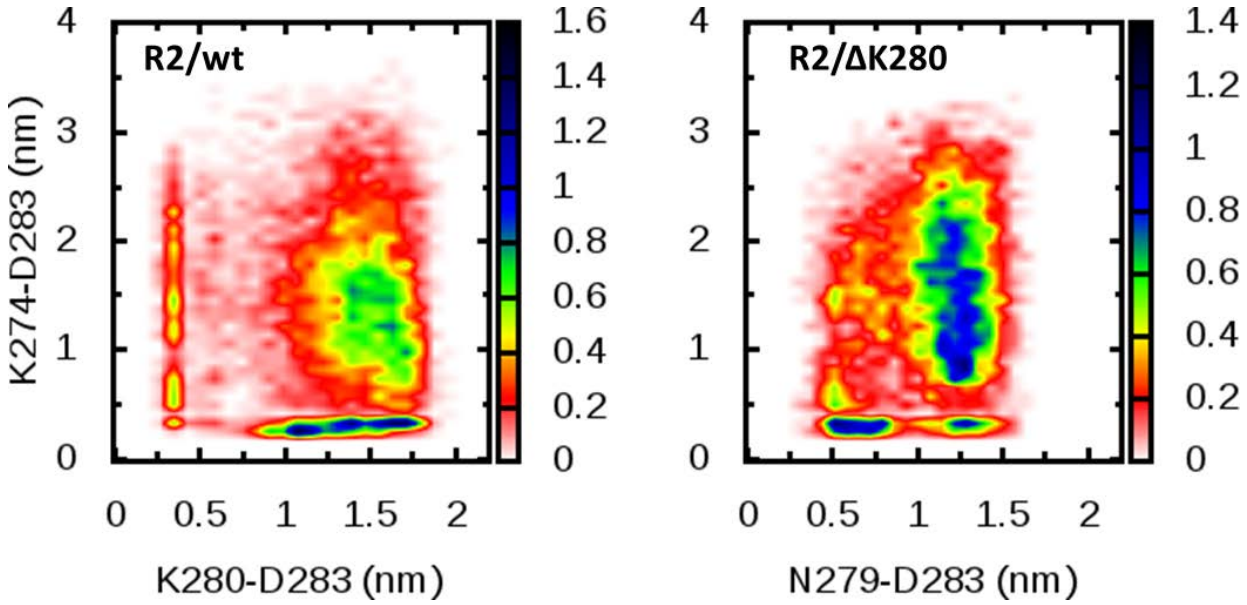


Figure S9 The probability of forming salt bridges for the wild-type R2/wt protein and its mutant ΔK280 R2/ΔK280 is evaluated according to the distance between residues with opposite charges. The only negative residue in the molecule is D283, which can form a salt bridge either with K274 or K280. For the mutant protein, N279 can, in principle, stabilize the molecule by forming a hydrogen bond with D283. Distances are defined between the center of mass of the COO group in the aspartic acid, and the center of mass of the group NH_3 (lysine) or NH_2 (asparagine).

This journal is © The Owner Societies 2013

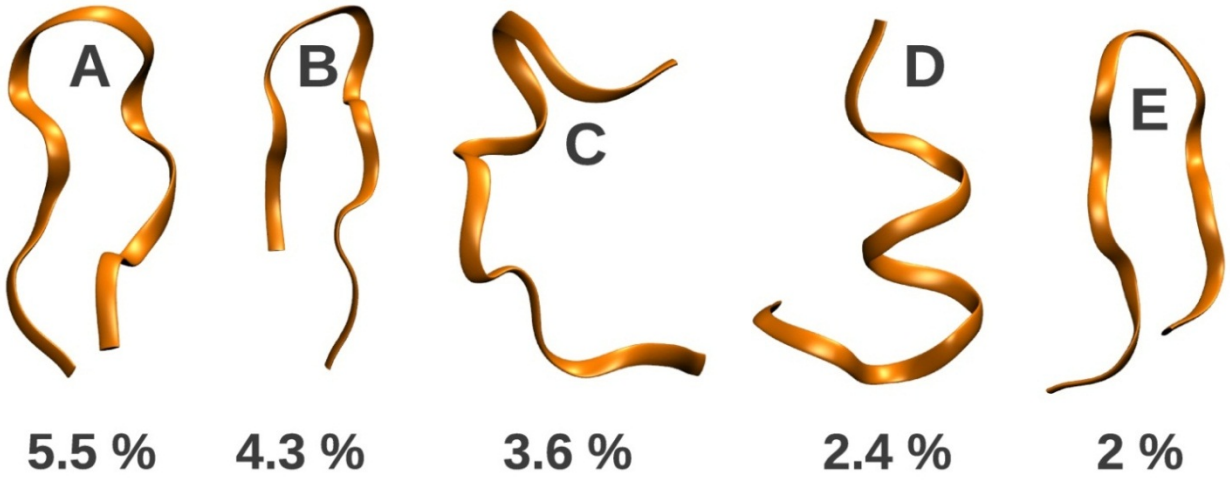


Figure S10 Snapshots showing the first five clusters of the wild-type R2/wt monomer in solution and their population.

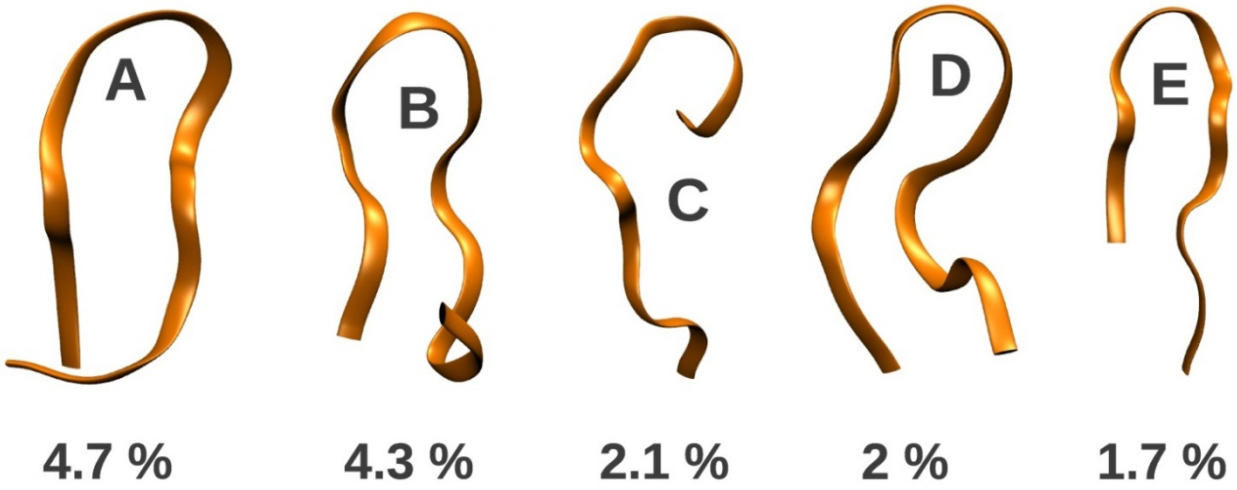


Figure S11 Population and representative structures of the first five clusters in solution for the mutant R2/ΔK280.

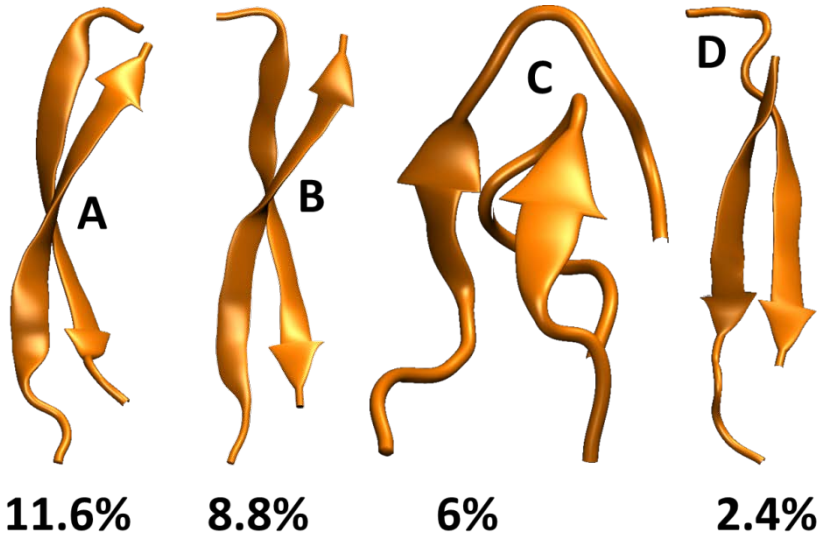


Figure S12 Representative conformations and their abundance for dimers of wild-type R2/wt (cluster A) and mutant R2/ΔK280(clusters B-D). Each cluster represents the most abundant conformation adopted by the same region in Fig. 8 c) and d). Region A is defined as R_g>1.04 and cos(θ)<−0.9; Region B is R_g>0.91 and cos(θ)<−0.55; Region C is 0.3<cos(θ)<0.54 and R_g<0.81; Region D is defined for cos(θ)>0.9.