Electronic Supplementary Information (ESI) Kinetic coherence underlies the dynamics of disordered proteins

Alexander Tenenbaum[a]

November 8, 2021

1 Pseudo-proteins

The primary sequences of Hip and of the three randomly scrambled pseudo-proteins are:

Hip: MRRRGEIDMATEGDVELELETETSGPERPPEKPRKHDSGAADLERVT-DYAEEKEIQSSNLETAMSVIGDRR SREQKAKQEREKELAKVTIKKEDLE-LIMTEMEISRAAAERSLREHMGNVVEALIALTN

P1h: RSVGLEIDEMDSMLESEDKEMVLVESKLIRSGGGNAVARRTEKK-TIKAKTIIGSNLRYPADREEDRDQRPA EAERRAPRAGERTHRTEDLPLEE-HEKMMLEEAEQLKTVQEEASALEEEIKAMTITNSV

P2h: PEELLRKILDERLSEEVVSTRGVSETDMEDVLNAMERLMNEIKAE-QPEPKEEERPTAKLTRAMLAHVGQAL IIRKDRDNAKYETAEAEMEEAESTRS-GDSTRKIKGTRAIMRHKRSGAEEIVQSDLEGE

P3h: MMSNNTIGRKVQMITDEKLEAASIIIEAEAEPDSKEETEILAML-GAKRLVDLKVLPAAVGEMKPESGRRRR EEVDTQTERREEEAEKLDQHATVERETRLDSSS-GIEYRRKEDMAARSHPLEKENETGL

The initial structures produced by the procedure of the I-TASSER platform and equilibrated in solvent at $T = 300$ K are shown in Figs. 1, 2, 3, and 4. At this temperature the Hia structure is open, while the three pseudo-proteins show a globular structure (P1A, P2A, P3A). During the 10 ns runs the pseudo-proteins' secondary structures' type and the extension have been slightly fluctuating, which is typical for transient structures of IDPs [1].

[1a] *Physics Department, Sapienza University, Piazzale Aldo Moro 5, 00185 Roma, Italy. E-mail: tenen314@gmail.com*

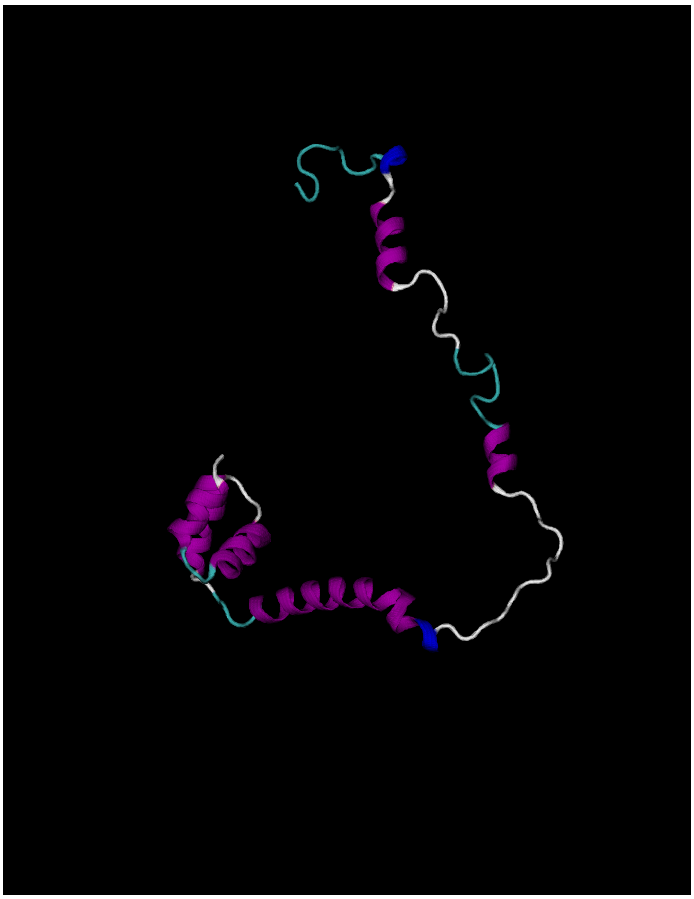

Figure 1: Huntingtin interactin protein at 300 K: HiA.

Figure 2: Pseudo-protein P1h at 300 K: P1A, globule.

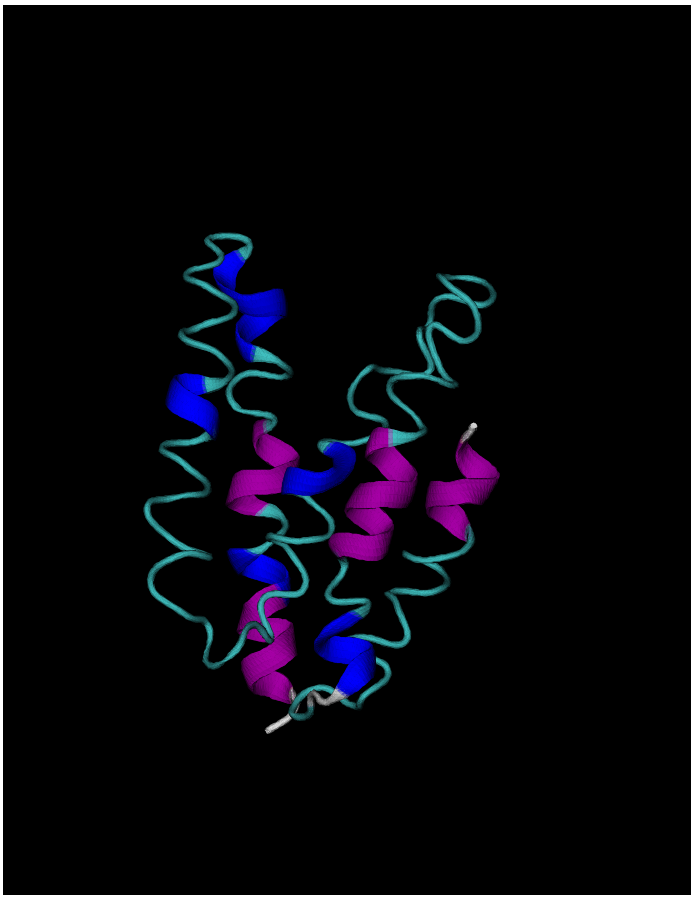

Figure 3: Pseudo-protein P2h at 300 K: P2A, globule.

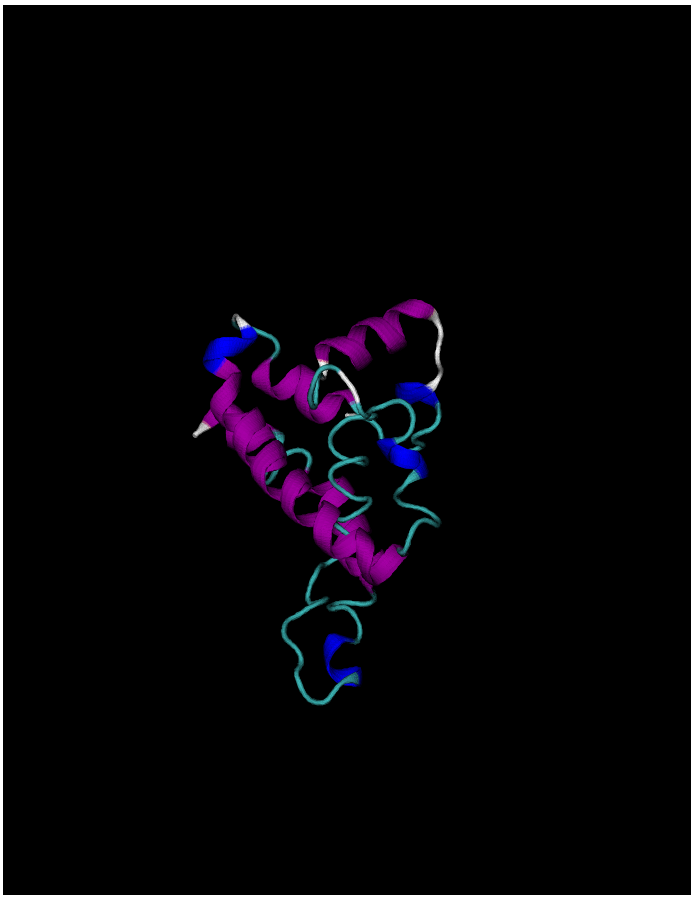

Figure 4: Pseudo-protein P3h at 300 K: P3A, globule.

The initial structures produced by the procedure of the I-TASSER platform, denaturated at $T = 1000$ K and equilibrated at $T = 370$ K, are shown in Figs. 5, 6, 7, and 8. At $T = 370$ K the four structures show a molten globular structure, consistent with their intrinsically disordered nature. Some short remnants of the secondary structures produced by the I-TASSER procedure are still present in all these structures.

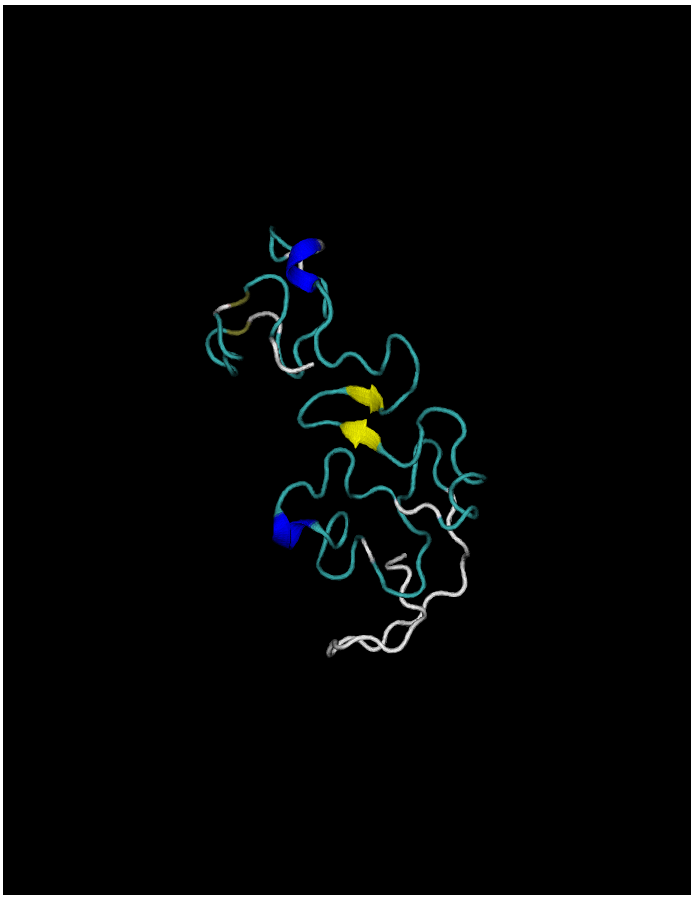

Figure 5: Huntingtin interacting protein at 370 K: HiC, molten globule.

2 Computer Simulation

In the simulation the short-range interactions were computed using a cut-off of 1.2 nm, and the long-range electrostatic interactions were computed using the particle-mesh Ewald sums (PME) [2] with a grid spacing of 0.16 nm. All H-bonds were constrained using the LINCS algorithm [3, 4]. The two 1 ns

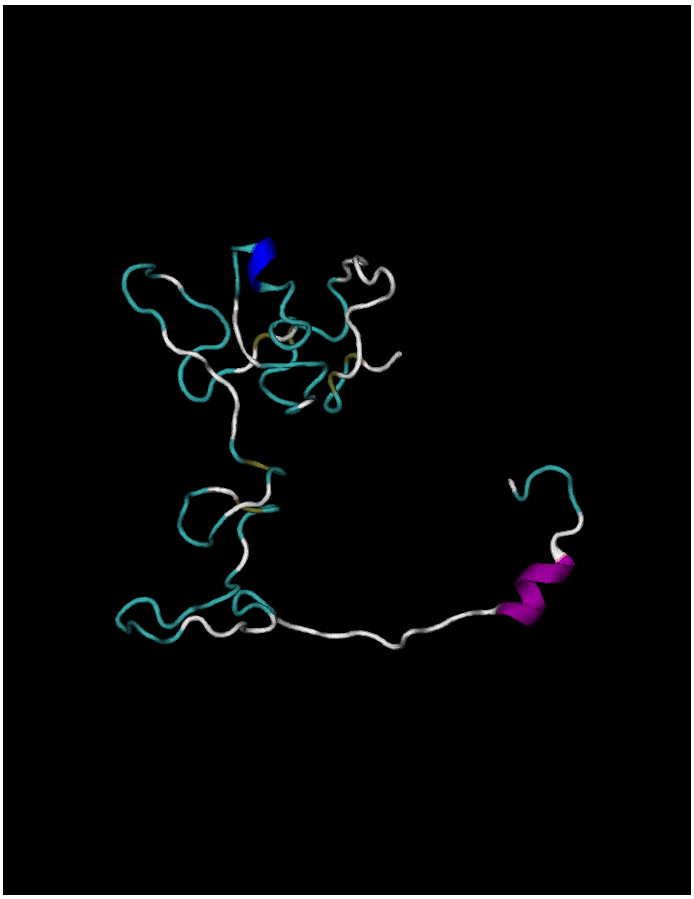

Figure 6: Pseudo-protein P1h at 370 K: P1C, molten globule.

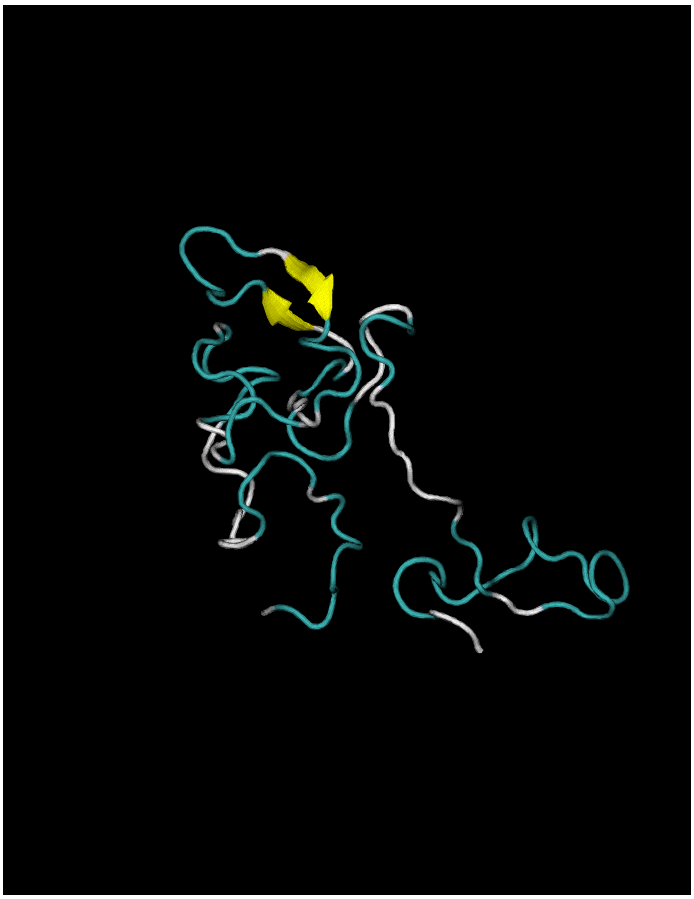

Figure 7: Pseudo-protein P2h at 370 K: P2C, molten globule.

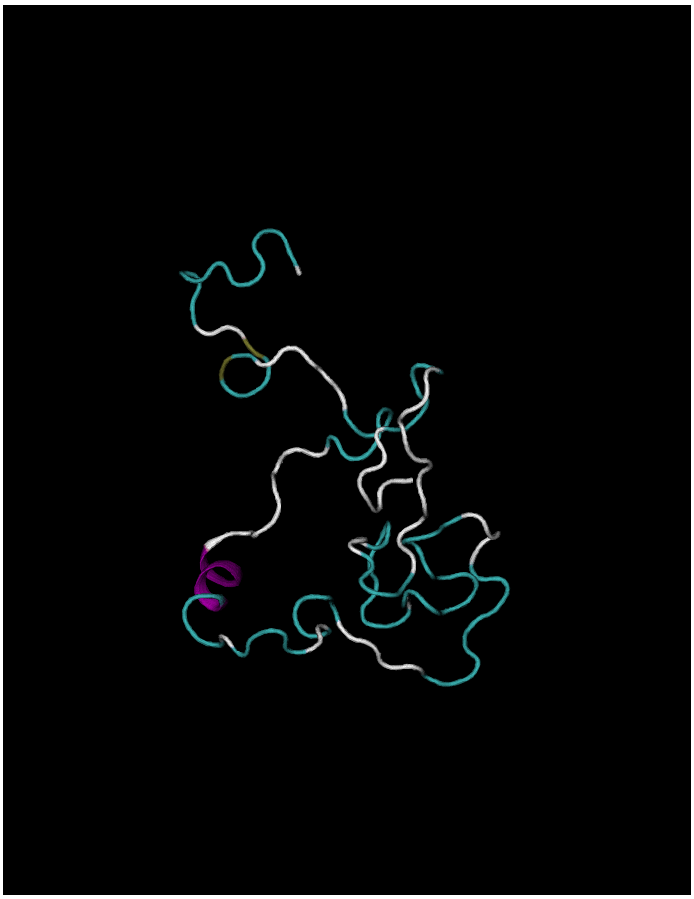

Figure 8: Pseudo-protein P3h at 370 K: P3C, molten globule.

Mol	*Residues*	*atoms*	*total*	*volume*	*bonds*
LyA	129	1960	33876	340.29	91
LyB	129	1960	33876	356.21	84
LyC	129	1962	29774	314.32	72
HiA	129	2044	58606	587.26	76
HiB	129	2044	58606	619.67	85
HiC	129	2044	56872	597.86	94
P1A	129	2044	21406	211.48	78
P1B	129	2044	21406	222.27	82
P1C	129	2044	51337	540.39	76
P2A	129	2044	24181	239.05	83
P2B	129	2044	24181	251.30	88
P2C	129	2044	34270	358.93	74
P3A	129	2044	20764	204.90	83
P3B	129	2044	20764	216.00	89
P3C	129	2044	20764	215.68	74
O1A	9	162	5179	52.01	4
O1B	9	162	5179	54.41	6
O2A	4	92	4460	44.58	0
O2B	4	92	4460	46.80	1

Table 1: 1st column: Molecule. 2nd column: Number of residues. 3rd column: Atoms of the solute. 4th column: Atoms of the whole system. 5th column: Volume of the box (nm^3). 6th column: Average number of intramolecular H-bonds

equilibration runs, performed after the energy minimization of the molecule in the solvent, were done by applying positional restraints of 1000 kJ mol^{-1} nm^{-1}. The 10 ns trajectories were computed in simple precision with a time step of 2.0 fs, in a NVT ensemble with a V-rescale temperature coupling [5]; these trajectories were computed eliminating the translation of the system. The 2 ps trajectories, over which the coherence times were measured, were computed in double precision with a time step of 0.2 fs, in a NVE ensemble without external coupling. No correction to the time evolution of the system was applied in these 2 ps stretches during which the $\Delta(t)$ distances were measured, in order not to perturb the dynamics. Further details of the computer simulation are reported in Table 1. -2

3 Coherence Times

The curves for P1h in Fig. 9 show that (a) approximately the first 30 $\tau_\pi^{(l)}$ are lower than those of the rest of eigenvectors oscillating with smaller amplitude;

(b) the $\tau_q^{(l)}$ and the $\tau_\pi^{(l)}$ do not vary significantly taking the system from P1A to P1B and to P1C. P1h's behaviour is thus similar to the behaviour of Hip, but very different from that of Lys. As for the other two pseudo-proteins, their curves of the PCs' coherent times are indistinguishable from those shown here in Fig. 9.

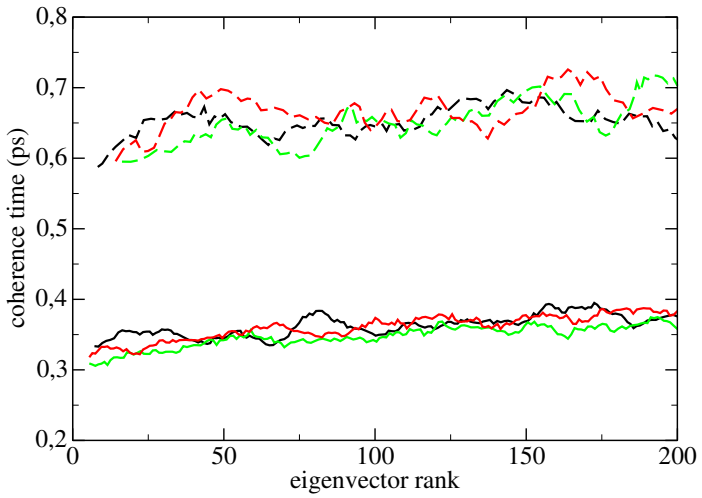

Figure 9: Coherence times of the eigenvectors in P1A (black lines), P1B (green lines), P1C (red lines). Positions: continuous lines; momenta: dashed lines. Every point is the average in a sliding window of 10 contiguous points.

References

[1] A. Battisti, G. Ciasca, A. Grottesi, A. Bianconi, and A. Tenenbaum, Temporary secondary structures in tau, an intrinsically disordered protein, *Molecular Simulation*, 2012, **38**, 525-533.

[2] T. Darden, D. York, and L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 1993, **98**, 10089–10092.

[3] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

[4] B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, *J. Chem. Theory Comput.*, 2008, **4**, 116–122.

[5] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.*, 2007, **126**, 8, 14101.