Conformational Diversity Induces Nanosecond-Timescale Chemical Disorder in the HIV-1 Protease Reaction Pathway

*Ana Rita Calixto, Maria João Ramos and Pedro Alexandrino Fernandes**

UCIBIO@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências Universidade do Porto, Rua do Campo Alegre s/n,4169-007 Porto, Portugal

SUPPORTING INFORMATION

The supporting information of this manuscript is presented below, and it has some important results that could help to understand the work presented above.

Methodology used in the initial enzyme modeling and MD simulations.

The initial system was constructed starting from the 4HVP X-ray structure downloaded from the Protein Data Bank (PDB). This initial structure contains the complete enzyme complexed with a substrate-based peptide inhibitor with the sequence Ac-Thr-Ile-Nle-[CH2-NH]-Nle-Gln-Arg-amide. This structure was modeled to a true substrate with the following sequence: Ac-Thr-Ile-

Met-[CO-NH]-Met-Gln-Arg-amide, in the same way as in a previous works[1-2] . A nucleophilic water molecule was added to the active center and the Asp25B carboxylate was protonated. It is experimentally known that this residue needs to be protonated to initiate the catalytic mechanism of this enzyme, and that its *p*Ka is high.

Prior to the MD simulation, 9960 TIP3P water molecules were added to the protein:substrate complex [3] , in a rectangular box of 88 Å x 67 Å x 71 Å. At least 12 Å were left between any atom of the complex protein-substrate surface and the external molecules of the solvent box. To equilibrate the modeled complex, we performed a first MD simulation without any restrictions on the structure. In this simulation, the catalytic water molecule diffused away from the active site to the solvent and the catalytic aspartates turned their side chains to each other (the well-known, very stable, "resting state" of HIV-1 PR). To circumvent this, we forced the protein to adopt the less abundant catalytic conformation by constraining the distance between the catalytic hydrogen atom of Asp25B and the carbonyl oxygen atom of the substrate. We used a harmonic potential having an equilibrium length of 1.80 Å between these two atoms and a force constant of 50 kcal.mol^{-1}Å^{-2}, using the same protocol as described in a previous work [1] . This procedure enriched the MD simulation in productive conformations, avoiding the need of a much longer MD simulation to collect a relevant number of them. The PME method [4] was used to calculate the Coulomb interactions with the real part truncated at 10 Å. Explicit van der Waals interactions were also truncated at 10 Å. We used the SHAKE algorithm [5] and a time step of 1 fs. In order to relax the system, removing possible tensions or clashes we started by a three-step minimization using Amber 12 simulation package[6] with parm 99SB force field. First, the water molecules were minimized with the remainder of the system fixed. In these calculations the steepest algorithm for 5000 cycles and conjugated gradient algorithm for the last 5000 steps. Then the

hydrogen atoms were minimized, fixing the remainder of the system (steepest descent algorithm to the first 5000 cycles, and conjugate gradient algorithm for the last 5000 steps). Finally, the positions of all atoms were minimized (steepest descent algorithm to the first 15000 steps and conjugate gradient algorithm for the last 15000 steps). Starting from the structures obtained after the minimization procedures, we ran MD simulations, with an initial warm-up of the system from 0 to 300 K during a 40 ps long simulation maintaining a constant volume and with periodic boundary conditions. Then we ran an MD equilibration on the whole system in the isothermal-isobaric ensemble (NPT) with the Langevin thermostat and isotropic position scaling, maintaining the temperature at 300 K and the pressure at 1 bar. Then the production dynamics ran during 130 ns with the same conditions.

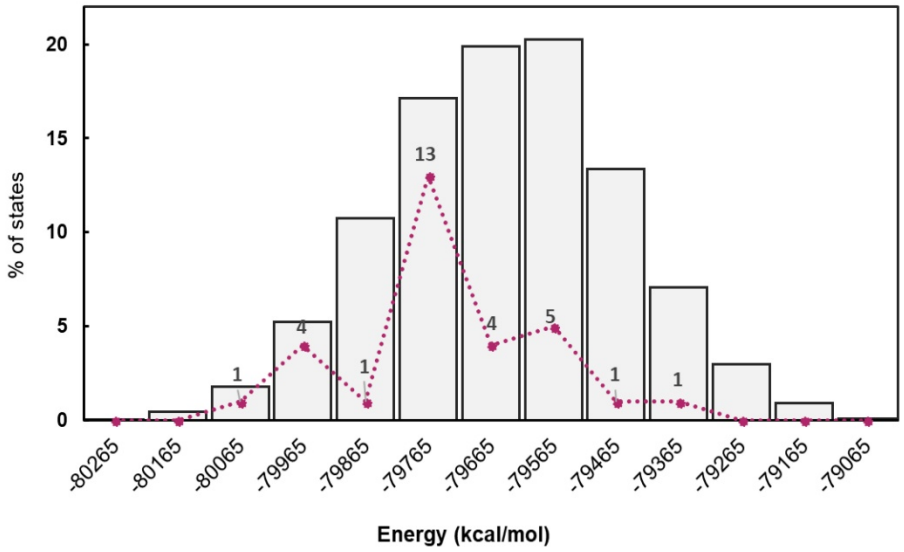

Figure S1. Energy distribution of the NPT ensemble generated by the MD calculation (grey bars) and of the microstates extracted for the QM/MM calculations (violet line). The distribution peaks around the average energy and decays slowly in both directions, very well in line with the expected NPT ensemble distribution of energy, which results from the product of the Boltzmann distribution, decaying very fast with increasing energy, and the density of states, increasing very fast with increasing energy. The states taken from the MD simulation to perform the QM/MM calculations correspond to well-populated regions of the MD ensemble, and follow the MD ensemble distribution reasonably well. As the phenomenological rate constant depends *linearly* in the state probability, but *exponentially* on the activation free energy, small differences in both QM/MM and MD probability distributions, like the ones shown above, do not have any significant impact in the resulting rate constant, when compared to the standard error of ~3 kcal/mol obtained in the activation free energies on DTM/MM calculations. The latter affects the rate constant by ~100 fold, up or down, which is orders of magnitude larger than the differences in distribution seen in the figure above.

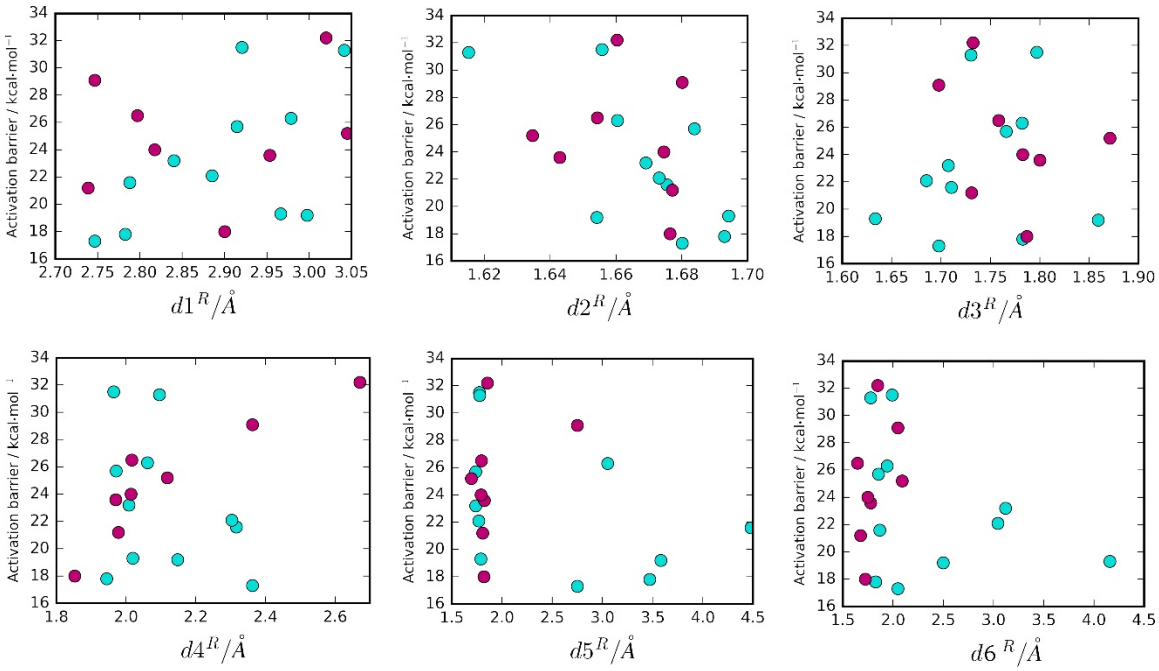

Figure S2. Correlations between six selected active site distances from reactant structures and the corresponding activation barriers. Purple and blue circles correspond to structures that react through the *One Asp* and *Two Asp* mechanism, respectively. There was no evident correlation between these distances and the barriers. However, for the *Two Asp* mechanism a small trend between *d1* and the activation barrier can be seen. The smallest distances were associated with small energies and large distances with large energies. The opposite trend can be observed for *d2*, and the same tendency for *d3*. The *One Asp* mechanism seemed to be more dependent on *d4*. Small values of *d4* and *d5* seemed to be required for this last mechanism.

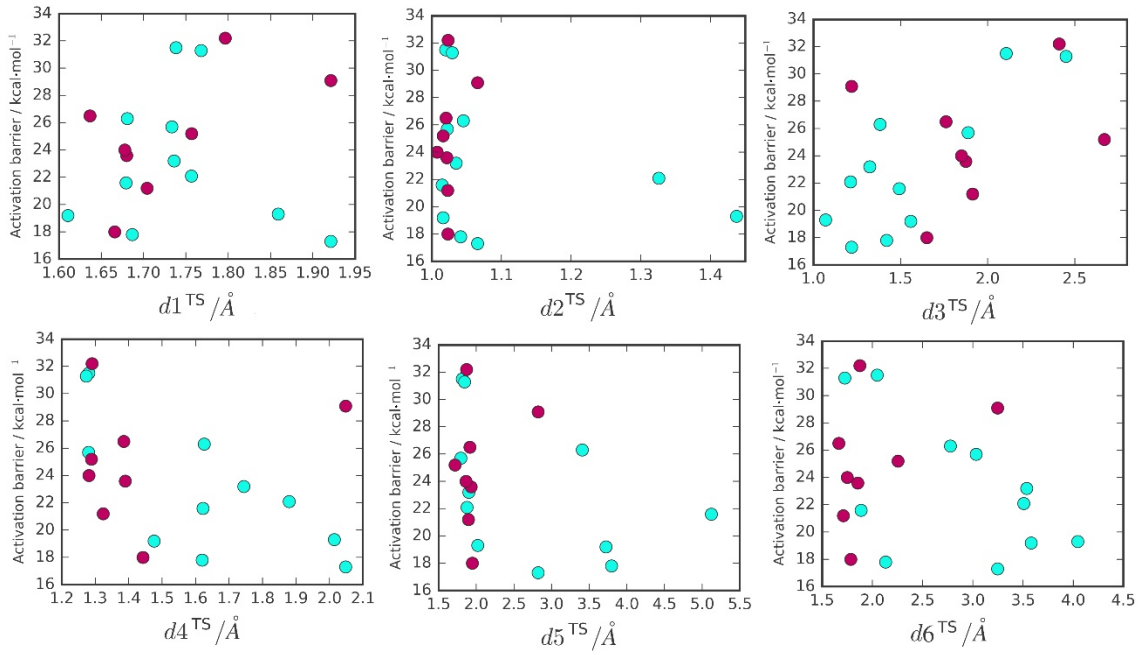

Figure S3. Correlations between six selected active site distances from transition state structures and the corresponding activation barriers. Purple and Blue circles correspond to structures that react through the *One Asp* and *Two Asp* mechanism, respectively. Once again, there was no evident correlation between these distances and the barriers. Small *d1* slightly correlate with the smallest energy for both mechanisms. The *Two Asp* Mechanism was associated with smaller values of *d3* and the *One Asp* mechanism with small values of *d4*, *d5* and *d6*.

Table S1. Specific reaction mechanism and activation barriers (Zero-point corrected Total Energy $\Delta E_0^{\ddagger}$, calculated at the M06-2X/6-311++G(2d,2p)-D3:ff99SB level of theory) for each selected reactant structure, respective imaginary frequency at the transition state and active site distances in both reactant and transition state structures.

				Reactant/Å						**TS/Å**					
Time (ns)	**Reaction mech.**	**Barrier/ kcal/mol**	**TS imaginary frequency (cm^{-1})**	***d*1**	***d*2**	***d*3**	***d*4**	***d*5**	***d*6**	***d*1**	***d*2**	***d*3**	***d*4**	***d*5**	***d*6**
100	*One Asp*	25.2	-679.4	3.05	1.63	1.87	3.05	1.70	2.09	1.76	1.02	2.67	1.29	1.72	2.26
101	*Two Asp*	19.3	-329.8	2.97	1.69	1.63	2.97	1.79	4.16	1.86	1.44	1.07	2.01	2.02	4.04
103	*One Asp*	23.6	-307.8	2.95	1.64	1.80	2.95	1.82	1.78	1.68	1.02	1.87	1.39	1.93	1.85
105	*Two Asp*	21.6	-234.4	2.79	1.68	1.71	2.79	4.48	1.87	1.68	1.01	1.49	1.62	5.12	1.89
107	*Two Asp*	26.3	-731.9	2.98	1.66	1.78	2.98	3.05	1.94	1.68	1.05	1.38	1.63	3.41	2.78
108	*One Asp*	21.2	-524.2	2.74	1.68	1.73	2.74	1.81	1.68	1.70	1.02	1.91	1.32	1.90	1.71
109	*Two Asp*	31.5	-769.6	2.92	1.66	1.80	2.92	1.77	1.99	1.74	1.02	2.11	1.28	1.81	2.05
110	*Two Asp*	17.8	-376.6	2.78	1.69	1.78	2.78	3.47	1.83	1.69	1.04	1.42	1.62	3.80	2.13
111	*One Asp*	18.0	-160.4	2.90	1.68	1.79	2.90	1.82	1.72	1.67	1.02	1.65	1.44	1.95	1.78
112	*One Asp*	29.1	-849.9	2.75	1.68	1.70	2.75	2.75	2.05	1.92	1.07	1.22	2.05	2.82	3.25
113	*Two Asp*	22.1	-358.0	2.89	1.67	1.69	2.89	1.77	3.04	1.76	1.33	1.21	1.88	1.88	3.51
116	*Two Asp*	19.2	-71.2	3.00	1.65	1.86	3.00	3.58	2.50	1.61	1.02	1.56	1.48	3.72	3.58
117	*Two Asp*	23.2	-710.9	2.84	1.67	1.71	2.84	1.74	3.12	1.74	1.03	1.32	1.74	1.90	3.54
118	*Two Asp*	25.7	-735.9	2.91	1.68	1.77	2.91	1.74	1.85	1.73	1.02	1.89	1.28	1.80	3.03
120	*Two Asp*	17.3	-923.9	2.75	1.68	1.70	2.75	2.75	2.05	1.92	1.07	1.22	2.05	2.82	3.25
121	*Two Asp*	31.3	-762.3	3.04	1.62	1.73	3.04	1.78	1.78	1.77	1.03	2.45	1.27	1.84	1.72
124	*One Asp*	26.5	-413.6	2.80	1.65	1.76	2.80	1.79	1.65	1.64	1.02	1.76	1.39	1.91	1.67
125	*One Asp*	24.0	-854.9	2.82	1.67	1.78	2.82	1.79	1.75	1.68	1.01	1.85	1.28	1.86	1.75
129	*One Asp*	32.2	-845.6	3.02	1.66	1.73	3.02	1.85	1.85	1.80	1.02	2.41	1.29	1.87	1.88

REFERENCES

1. Ribeiro, A. n. J.; Santos-Martins, D.; Russo, N.; Ramos, M. J.; Fernandes, P. A., Enzymatic Flexibility and Reaction Rate: A QM/MM Study of HIV-1 Protease. *ACS Catal*. **2015,** *5* (9), 5617-5626.

2. Calixto, A. R.; Ramos, M. J.; Fernandes, P. A., Influence of Frozen Residues on the Exploration of the PES of Enzyme Reaction Mechanisms. *J. Chem. Theory Comput*. **2017,** *13* (11), 5486-5495.

3. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **1983,** *79* (2), 926-935.

4. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Chem. Phys*. **1995,** *103* (19), 8577-8593.

5. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*. **1977,** *23* (3), 327-341.

6. Case, D.; Darden, T.; Cheatham III, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K., AMBER 12; University of California: San Francisco, 2012. **2010**, 1-826.