

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

### Unfolding Dynamics of a Photo-Switchable Helical Peptide

|                               |   |
|-------------------------------|---|
| Journal:                      | <i>The Journal of Physical Chemistry</i>  |
| Manuscript ID                 | jp-2020-04017p.R1   |
| Manuscript Type:              | Article   |
| Date Submitted by the Author: | n/a   |
| Complete List of Authors:     | Auvray, Francois; University of Nottingham, School of Chemistry<br>Hirst, Jonathan; University of Nottingham, School of Chemistry |
|                               |   |

SCHOLARONE™  
Manuscripts

# Unfolding dynamics of a photo-switchable helical peptide

Francois Auvray and Jonathan D. Hirst\*

*School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD,  
U.K.*

E-mail: jonathan.hirst@nottingham.ac.uk

Phone: +44 115 951 3478. Fax: +44 115 951 3562

## Abstract

We present an atomistic force field for the azo-moiety of the photo-switchable FK-11-X peptide. We use the parameters to study the unfolding of the peptide through molecular dynamics simulations. The unfolded ensemble contains many different structures, ranging from a partially unfolded peptide to a fully unfolded structure. The averaged computed far-ultraviolet circular dichroism (CD) spectrum of the set of structures, which were simulated using the newly developed force field, agrees well with experiment. The rate of the simulated unfolding process was estimated to have a time constant of  $5.80 \pm 0.03$  ns from the time-evolution of the CD spectrum of the peptide, computed from the backbone conformations sampled over 40 simulated trajectories. Our estimated time constant is faster than, but not inconsistent with, previous experimental estimates from time-resolved infrared and optical rotatory dispersion spectroscopy.

## Introduction

The cooperative nature of protein folding suggests that individual  $\alpha$ -helices should not fold in isolation. However, Brown and Klee<sup>1</sup> observed, using circular dichroism (CD) spectroscopy, partial helical structure in the isolated C-peptide of ribonuclease A. Subsequently, the thermodynamics and kinetics of the folding of  $\alpha$ -helices have been widely studied<sup>2</sup> and these studies have elucidated many important aspects of protein folding. One example of many was examination of the stabilizing thermodynamic influence of side-chain interactions in isolated helices in the absence of the complexities of a surrounding protein environment.<sup>3</sup>

The unfolding dynamics of helical peptides have been studied using temperature-jump (T-jump) techniques in conjunction with various spectroscopies. T-jump experiments move a system away from equilibrium by rapid heating of the solvent by several degrees within a few nanoseconds by irradiating the sample with a pulsed infrared laser. They provide insights on microsecond and nanosecond timescales, which are of particular interest for early events in protein folding. Techniques employed in these studies include infrared spectroscopy<sup>4</sup> and ultraviolet (UV) resonance Raman spectroscopy.<sup>5</sup> Khuc et al.<sup>6</sup> used a T-jump far-UV CD experiment with a time step of 12 ns to follow the loss of helical structure upon thermal denaturation of a poly(glutamic acid) peptide, exploiting a well-known empirical relationship between the intensity of a distinctive negative band at 222 nm and the  $\alpha$ -helical content.<sup>7</sup> The influence of the backbone conformation on far-UV CD spectra has been explored theoretically over many years,<sup>8,9</sup> including the role of main-chain hydrogen bond lengths<sup>10</sup> and the sensitivity to the backbone dihedral angles.<sup>11</sup> This provides a direct connection between atomistic models of peptide (un)folding and CD spectroscopy.

Molecular dynamics (MD) simulations on nanosecond and microsecond timescales complement experimental investigations and furnish atomistic detail about the folding mechanism and the associated folding landscape, including the transiently stable conformations.<sup>12,13</sup> Simulations of peptides are computationally more tractable than simulations of proteins, because of their smaller size. However, this is somewhat offset by their greater conformational

1  
2  
3 flexibility, which presents a non-trivial sampling challenge. Another key aspect is the quality  
4 of the empirical force fields used in the simulations. Confidence in the sampling and the  
5 force field can be developed by computing spectroscopic observables, with the caveat that  
6 reproduction of experiment is a necessary but not sufficient condition for establishing the  
7 adequacy of the simulation methodology.  
8  
9

10  
11  
12  
13 There are several approaches to computing CD spectra from the atomic coordinates  
14 from an MD simulation (or indeed an experimentally determined structure) of a peptide or  
15 protein. These have recently been reviewed in detail elsewhere.<sup>14</sup> Empirical approaches, such  
16 as BeStSel,<sup>15</sup> PDB2CD<sup>16</sup> and SESCA,<sup>17</sup> take various structural descriptors derived from the  
17 atomic coordinates and some “basis” spectra to generate a computed spectrum. Quantum  
18 chemical methods can be used. Time-dependent density functional theory (TDDFT) has  
19 been employed<sup>18</sup> and inventive approaches applicable to large molecular systems have been  
20 proposed,<sup>19</sup> involving simplified Tamm-Dancoff density functional theory.  
21  
22  
23  
24  
25  
26  
27  
28

29 Calculations based on exciton theory,<sup>20,21</sup> with an effective Hamiltonian, are more approx-  
30 imate than fully quantum mechanical methods. They are applicable to large biomolecules  
31 and the first principles nature of the approach means it should be more applicable than  
32 fully empirical methods to partially unfolded conformations that may not be present in  
33 structural databases. Local excitation properties and exciton couplings can be computed  
34 with TDDFT.<sup>22</sup> Much work has also been done on developing parameters from higher level  
35 multi-configurational ab initio methods for the relevant electronic excitations of the peptide  
36 backbone.<sup>11</sup> These parameters and the computational methodology are available via the  
37 DichroCalc web interface<sup>23</sup> and a github repository.<sup>14</sup>  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 Simulation of conformational dynamics and the associated CD spectra is proving to be  
48 a useful strategy for studying peptide and protein folding. A combination of CD spec-  
49 troscopy, exciton calculations, standard MD simulations, and advanced-sampling methods  
50 allowed Hildebrand et al.<sup>24</sup> to pinpoint a destabilization effect of silica on specific structural  
51 motifs of chymotrypsin. Using a similar strategy, Michaelis et al.<sup>25</sup> demonstrated that an  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 appropriate sampling of structures between a folded helix and its unfolded state permits one  
4 to generate reliable computed spectra of transient states. First and Webb<sup>26</sup> performed MD  
5 folding simulations of a highly charged peptide in different solvents and bound to the sur-  
6 face of a methyl-terminated self-assembled monolayer. They found quantitative agreement  
7 between previous CD experiments and their MD simulations of solvent- and surface-induced  
8 conformational changes of the peptide. A related approach was used by Gattuso et al.<sup>27</sup> to  
9 study the light-triggered conformational change of an  $\alpha$ -helical peptide of 27 amino acids  
10 with a retinal-like photoswitch covalently linked to the backbone. They correctly predicted  
11 the reduction in helicity of the peptide upon E to Z photoisomerisation of the switch.  
12  
13  
14  
15  
16  
17  
18  
19  
20

21 Photoisomerisation of a molecular switch attached to the peptide backbone is strategy  
22 often employed to access fast timescales.<sup>28</sup> Different molecular photo-switches have been  
23 studied. Of these, the azobenzene moiety has been frequently utilised.<sup>29,30</sup> Azobenzene can  
24 be isomerised from the *trans* to *cis* and *cis* to *trans* conformation using a 370 nm and a  
25 visible light source (460 nm), respectively.<sup>31</sup> Several groups have studied the dynamics of  
26 this photo-induced reaction and the isomerisation of the azobenzene appears to occur within  
27 one picosecond.<sup>32-35</sup> The changes in the molecular geometry that occur over its isomerisation  
28 can be used to induce an imbalance in the structure of larger molecules, such as peptides  
29 and proteins.<sup>36-38</sup>  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 The study that we present here focuses on the azobenzene cross-linked 16 amino-acid  
40 peptide FK-11-X.<sup>29,39</sup> The isomerisation of the azobenzene triggers the unfolding/folding  
41 process of the peptide. Experimental studies agree that, following the *trans* to *cis* azoben-  
42 zene isomerisation, peptide conformational changes related to the unfolding process occur  
43 on the 100 ns to 1  $\mu$ s time scale.<sup>40,41</sup> Xia *et al.*<sup>42</sup> used a combination of quantum mechan-  
44 ics/molecular mechanics (QM/MM) methods with a classical MM atomistic force field to  
45 simulate the unfolding process of a peptide with a similar sequence to FK-11-X. Although  
46 this approach successfully simulated the unfolding process of the peptide, the timescale of  
47 the dynamics did not match with the values found experimentally. In our study, using a  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 purely classical MD approach, we sample significantly more of the conformational space.  
4

5 In this study, we develop an atomistic force field for the azo-moiety of the FK-11-X  
6 peptide for use in MD simulations. We will describe the methodology we employed and the  
7 set of parameters we obtained. We utilise this atomistic force field to run 40 MD trajectories  
8 simulating the unfolding process of FK-11-X. We characterise the folding state of the peptide  
9 over time using various structural properties and we compare the trajectories with each other.  
10 We calculate the rate of unfolding and compare this to experimental data. Using ensembles  
11 of computed molecular geometries, we calculate the time-evolution of the CD spectrum of  
12 FK-11-X as it unfolds. We connect these spectra to structural aspects of FK-11-X. Our  
13 results confirm that MD simulations can complement experimental measurements to give  
14 insights into transient conformations of small peptides as they unfold. Our assessment of the  
15 rate of the simulated unfolding process corresponds well to the value found experimentally,  
16 giving some confidence to the new atomistic force field parameters presented herein.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

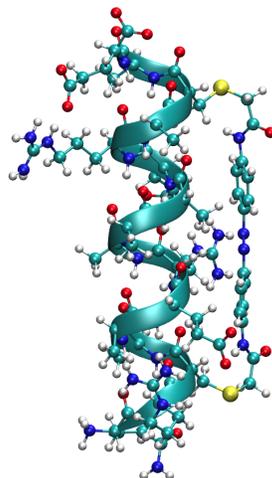
## 31 **Methods**

32

33 We built the peptide and the azo-moiety with the Avogadro software<sup>43</sup> (version 1.2.0), using  
34 the feature that allows one to generate an  $\alpha$ -helix structure from an amino acid sequence. We  
35 constructed the azo-moiety in its *trans* configuration, placed it close to the position where it  
36 should be when it is linked to the peptide and created the bonds between the azo-moiety and  
37 the two sulfur atoms of the cysteines. The resulting structure after minimisation is shown  
38 in Figure 1. All the molecular visualisations presented in this paper were rendered using the  
39 Tachyon extension in VMD.<sup>44</sup>  
40  
41  
42  
43  
44  
45  
46

47 We solvated the molecule using a box of TIP3P water.<sup>45</sup> We determined the appropriate  
48 size of the water box from the length of the peptide and the cutoff distance of the non-  
49 bonded interaction we used for the simulation. We performed a preliminary MD simulation  
50 with the unfolded peptide to estimate of the maximum length of the peptide. The distance  
51 between the C- and N-termini never exceeded 35 Å. The cutoff distance for the non-bonded  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 interactions is 12 Å, so we set the edge of the cubic water box at 50 Å. The solvated system  
4  
5 comprises 15570 atoms.  
6  
7

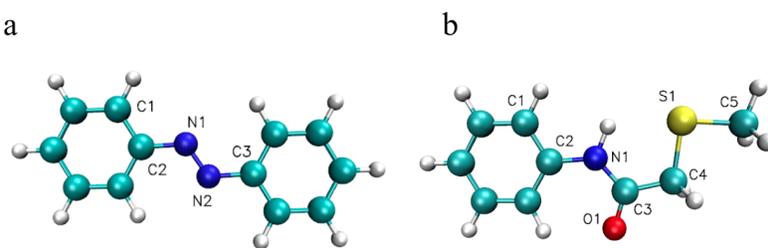


8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26 Figure 1: Structure of the FK-11-X peptide. Its amino acid sequence is Ac-Glu-Ala-Cys<sup>AZO</sup>-  
27 Ala-Arg-Glu-Ala-Ala-Ala-Arg-Glu-Ala-Ala-Cys<sup>AZO</sup>-Arg-Gln-NH<sub>2</sub>.  
28  
29

30 We performed the simulation using NAMD.<sup>46</sup> It models the bonded and non-bonded  
31 interactions of each atom with all the other atoms of the system with a potential energy  
32 function whose form is well established in the field. We used the CHARMM36 All-Hydrogen  
33 force field<sup>47–49</sup> to model the peptide. For the azo-moiety, Carstens<sup>50</sup> developed parameters  
34 that have been used in some studies;<sup>39,40</sup> and Böckmann *et al.*<sup>51</sup> proposed a set of force  
35 field parameters for an azobenzene derivative. However, even if we could use these for  
36 the azobenzene, we would still need parameters for the atoms that link the peptide to  
37 the azobenzene. We used the online interface<sup>52</sup> of the CHARMM General Force Field<sup>53</sup>  
38 (CGenFF) to determine the required parameters by analogy with existing parameters of  
39 small organic molecules. This tool provides a penalty score for each attributed parameter.  
40 If this score is higher than 50, it is recommended to develop a customized force field for the  
41 corresponding bonds/angles. In the case of the azo-moiety, this score remains below 50 except  
42 for the parameters of dihedral angles. These dihedral angles have a relatively low contribution  
43 to the total potential energy. So we performed some preliminary tests using the CHARMM36  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 All-Hydrogen force field for the peptide, the parameters proposed by Böckmann *et al.*<sup>51</sup> for  
4 the azobenzene and the parameters generated by CGenFF for the atoms connecting the  
5 azobenzene and the peptide. However, the FK-11-X did not behave as expected. The  
6 strain applied by the *cis* azobenzene did not induce a decrease in the helical content of  
7 the peptide structure even after 25 ns of simulation. Time-resolved IR spectroscopy<sup>41,54</sup> and  
8 optical rotatory dispersion spectroscopy<sup>40</sup> (ORD) experiments shown that the first changes in  
9 conformation appear within a few nanoseconds after the isomerisation of the azobenzene. We  
10 concluded that the force field parameters given by CGenFF were not adequate. We developed  
11 a force field for the whole molecule, rather than use Böckmann *et al.*'s parameters<sup>51</sup> for the  
12 azobenzene, to ensure internal consistency.

13  
14  
15 We followed the procedure proposed by Vanommeslaeghe *et al.*<sup>53</sup> It consists of optimising  
16 the parameters to obtain similar potential energy profiles calculated from geometries mini-  
17 mized with QM and MM approaches. The geometries and energy profiles from QM are the  
18 target of the optimisation process. We considered the azo-moiety as comprising two smaller  
19 molecules to reduce the number of degrees of freedom of the system. We optimized separately  
20 the parameters of the azobenzene (Figure 2a) and the atom chain that links the azobenzene  
21 to the peptide (Figure 2b). We added a sulfur and a methyl to the carbon C<sub>4</sub> to determine  
22 the parameters of the link between the cysteines of the peptide and the azo-moiety.



48 Figure 2: **a.** Azobenzene **b.** Structures used for the force field parameter optimisation of  
49 the linker between the azobenzene and the peptide.

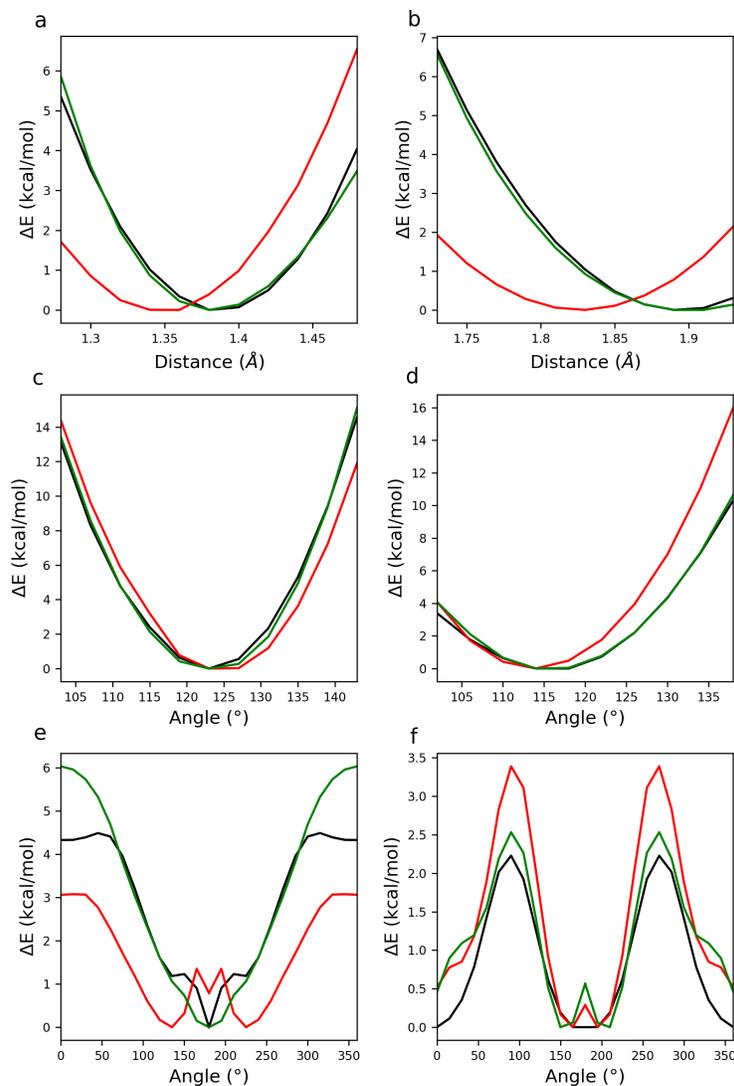
50  
51  
52 The first step of the procedure is the partial atomic charge determination to attribute  
53 a charge to each atom. We generated the initial guesses with CGenFF. The penalty scores  
54 were below ten for all the atoms except for four atoms of the azobenzene (C<sub>2</sub>, N<sub>1</sub>, N<sub>2</sub> and  
55  
56  
57  
58  
59  
60

1  
2  
3 C<sub>3</sub>). However, for these four atoms, differences between the charges generated by CGenFF  
4 and the charges proposed by Böckmann *et al.*<sup>51</sup> were below 5%. So we kept the initial  
5 guess of the partial atomic charges generated by CGenFF for the optimisation and for the  
6 simulations. We computed the QM geometries with Q-Chem<sup>55</sup> at the MP2/6-31G\* level of  
7 theory. We calculated the energy profile for all the bonds, the angles and the dihedral angles  
8 involving atoms that are indexed in Figure 2. The parameters of the benzene rings and the  
9 parameters of the bonds/angles involving hydrogens are generated by CGenFF.  
10  
11  
12  
13  
14  
15  
16

17 We performed the MM minimisation of the structures using NAMD. We adjusted the  
18 parameters ‘by hand’ until the energy profiles from MM fit with the profile from QM. We  
19 first optimised the parameters of the bonds, then the parameters of the angles and finally  
20 the parameters of the dihedral angles. Examples of the potential energy surfaces are shown  
21 in Figure 3. Both the MM geometries computed with initial guess and optimised param-  
22 eters are represented. The energy profiles calculated with the initial guess from CGenFF are  
23 significantly different compared to the profiles based on QM calculations. These differences  
24 may be responsible for the qualitatively incorrect behaviour we observed in the preliminary  
25 simulations using the initial guess. The profiles obtained with the optimised parameters are  
26 closer to the target and should improve the fidelity of the simulation. The optimised param-  
27 eters are presented in Tables 1 and 2, and are compared in the Supplementary Information  
28 (Tables S1 and S2).  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 The MD simulations were performed with NAMD 2.12 with a 2 fs timestep. The tempera-  
42 ture was maintained at 295 K using Langevin dynamics parameters with a friction coefficient  
43 of 5 ps<sup>-1</sup>. The pressure was kept constant at 1 atm with a Langevin piston that modifies  
44 the size of the unit cell to adjust the pressure. The maximum distance of non-bonded inter-  
45 actions was 12 Å. We used cubic periodic boundary conditions. Electrostatics were treated  
46 using the particle-mesh Ewald method<sup>56</sup> with a grid spacing of 1 Å. We used the SETTLE  
47 algorithm<sup>57</sup> and the ShakeH algorithm<sup>58</sup> to constrain water molecules and the other bonds  
48 involving hydrogen atoms, respectively. We performed two sets of simulations. The first  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 is composed of 40 trajectories where the azobenzene is isomerised in the *cis* form and a  
4 second set of 10 trajectories with the *trans* configuration of the azobenzene, which served as  
5 a control. Each of the trajectories consisted of 120 ns of production dynamics. In each case,  
6 the initial structure was minimized for 200 steps and was equilibrated for 200 ps, which was  
7 sufficient based on the convergence of the RMSD (Supplementary Information, Figure S1).  
8  
9  
10  
11  
12  
13



49 Figure 3: Examples of potential energy profiles of the structure shown in Figure 2b. The  
50 profile from MP2/6-31G\* geometries are plotted in black. The profile from MM geometries  
51 minimised using initial guess and optimised parameters are plotted in red and green, respec-  
52 tively. **a.** N<sub>1</sub>C<sub>3</sub> bond **b.** C<sub>4</sub>S<sub>1</sub> bond **c.** C<sub>1</sub>C<sub>2</sub>N<sub>1</sub> angle **d.** C<sub>3</sub>C<sub>4</sub>S<sub>1</sub> angle **e.** O<sub>1</sub>C<sub>3</sub>C<sub>4</sub>S<sub>1</sub> dihedral  
53 angle **f.** C<sub>1</sub>C<sub>2</sub>N<sub>1</sub>C<sub>3</sub> dihedral angle.  
54  
55  
56  
57  
58  
59  
60

Table 1: Optimised force field parameters of the azobenzene (Figure 2a).

| <b>Dihedral angle</b>                                       | $k$ (kcal/mol) | $n$ | $\delta$ ( $^\circ$ ) | <b>Angle</b>                                 | $k$ (kcal/mol) | $\theta_0$ ( $^\circ$ ) | <b>Bond</b>                   | $k$ (kcal/mol) | $b_0$ ( $\text{\AA}$ ) |
|---|----------------|-----|-----------------------|--|----------------|-------------------------|-------------------------------|----------------|------------------------|
| C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 2.5            | 2   | 180                   | C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> | 50             | 116.5                   | C <sub>2</sub> N <sub>1</sub> | 300            | 1.415                  |
| C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> C <sub>3</sub> | 23.0           | 2   | 180                   | C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 75             | 106.0                   | N <sub>1</sub> N <sub>2</sub> | 500            | 1.295                  |

Table 2: Optimised force field parameters of the molecule shown in Figure 2b.

| <b>Dihedral angle</b>                                       | $k$ (kcal/mol) | $n$ | $\delta$ ( $^\circ$ ) | <b>Angle</b>                                 | $k$ (kcal/mol) | $\theta_0$ ( $^\circ$ ) | <b>Bond</b>                   | $k$ (kcal/mol) | $b_0$ ( $\text{\AA}$ ) |
|---|----------------|-----|-----------------------|--|----------------|-------------------------|-------------------------------|----------------|------------------------|
| C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> C <sub>3</sub> | 1.20           | 2   | 180                   | C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> | 70             | 115.2                   | C <sub>2</sub> N <sub>1</sub> | 380            | 1.420                  |
| C <sub>2</sub> N <sub>1</sub> C <sub>3</sub> O <sub>1</sub> | 2.65           | 2   | 180                   | C <sub>2</sub> N <sub>1</sub> C <sub>3</sub> | 60             | 120.8                   | N <sub>1</sub> C <sub>3</sub> | 460            | 1.385                  |
| C <sub>2</sub> N <sub>1</sub> C <sub>3</sub> C <sub>4</sub> | 2.00           | 2   | 180                   | N <sub>1</sub> C <sub>3</sub> O <sub>1</sub> | 92             | 124.5                   | C <sub>3</sub> O <sub>1</sub> | 830            | 1.275                  |
| N <sub>1</sub> C <sub>3</sub> C <sub>4</sub> S <sub>1</sub> | 0.50           | 1   | 180                   | N <sub>1</sub> C <sub>3</sub> C <sub>4</sub> | 68             | 118.0                   | C <sub>3</sub> C <sub>4</sub> | 250            | 1.542                  |
| O <sub>1</sub> C <sub>3</sub> C <sub>4</sub> S <sub>1</sub> | 0.50           | 1   | 0                     | O <sub>1</sub> C <sub>3</sub> C <sub>4</sub> | 78             | 122.0                   | C <sub>4</sub> S <sub>1</sub> | 225            | 1.893                  |
| C <sub>3</sub> C <sub>4</sub> S <sub>1</sub> C <sub>5</sub> | 0.24           | 1   | 180                   | C <sub>3</sub> C <sub>4</sub> S <sub>1</sub> | 40             | 115.3                   | S <sub>1</sub> C <sub>5</sub> | 225            | 1.893                  |
|   |                |     |                       | C <sub>4</sub> S <sub>1</sub> C <sub>5</sub> | 58             | 99.0                    |                               |                |                        |

There are two methods to trigger the isomerisation of the azo-moiety. One can use a QM approach<sup>42</sup> or a potential energy triggering.<sup>59</sup> We used the second option. The principle of this method is to change the parameters of the dihedral angle responsible of the azobenzene configuration (C-N-N-C). We modified the multiplicity of this dihedral from two to one (see Figure 4a). This modification changes the number of minima from two to one. The unique energy minimum is located at  $0^\circ$ , which corresponds to the *cis* configuration of the azobenzene. We performed a 1.5 ps simulation with this modification for each structure starting from the conformation of the last step from the equilibration trajectory. The averaged value of the C-N-N-C angle of the azobenzene over time is plotted in Figure 4b. We see that the dihedral angle decreases from  $180^\circ$  to approximately  $60^\circ$  before the first picosecond. The angle passes the potential barriers located at  $-90^\circ$  and  $+90^\circ$  when  $n = 2$  and so the azobenzene will remain in the *cis* conformation even with  $n = 2$ . We used the conformation of the last step from the 40 isomerisation trajectories for the 120 ns simulations. For these simulations, the multiplicity parameter of the C-N-N-C dihedral angle was  $n = 2$ .

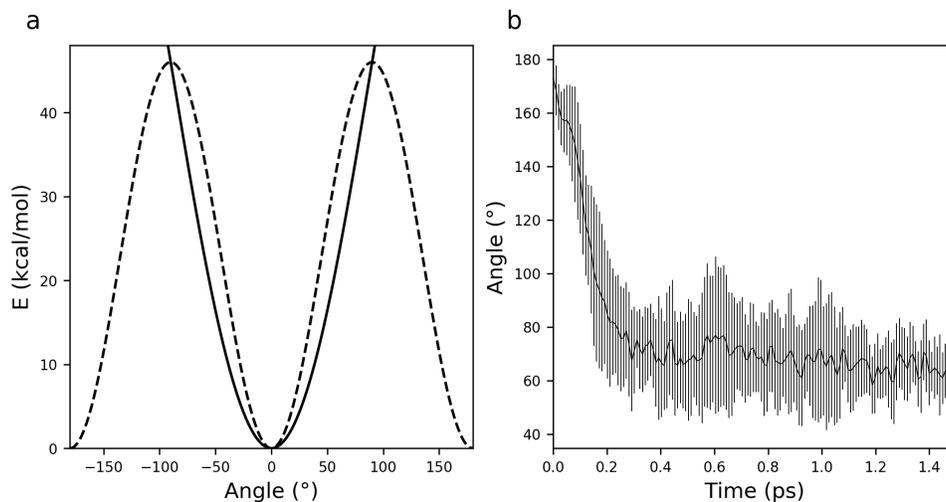


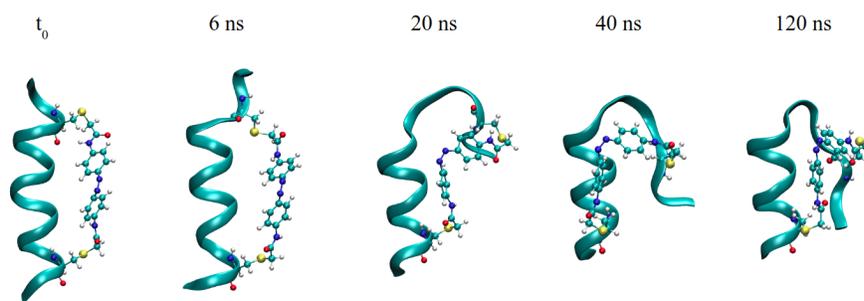
Figure 4: **a.** C-N-N-C dihedral angle energy function with multiplicity  $n = 1$  (solid line) and  $n = 2$  (dashed line). **b.** Averaged C-N-N-C dihedral angle value over the isomerisation trajectories. The error bars correspond to one standard deviation.

Theoretical CD spectra were calculated using the DichroCalc web interface<sup>23</sup> for individual frames of the simulated trajectories. DichroCalc constructs an exciton Hamiltonian matrix using electronic excitations of the local peptide chromophores. The Hamiltonian models how electronic transitions interact with one another. The sign and magnitude of the interactions depend on the relative orientation and separation of chromophores, which in turn arise from the precise structure of the peptide. For each peptide chromophore two transitions,  $n \rightarrow \pi^*$  and  $\pi_{nb} \rightarrow \pi^*$ , were considered using an ab initio parameter set.<sup>11</sup> Other transitions, such as the  $\pi_b \rightarrow \pi^*$  and charge transfer transitions, were not included. The calculated line spectra were convoluted with Gaussian bands of bandwidth 12.5 nm.

## Results & Discussion

The trajectories are, on the whole, relatively similar to one another. The similarity amongst the trajectories was assessed by computing the root mean squared inner product between pairs of trajectories,<sup>60</sup> where each trajectory was represented by the first principal component of the backbone dynamics computed over the final 5 ns. The tenth trajectory appears to be the most representative, as it has the highest average RMSIP value (0.51) over the 39

1  
2  
3 pairwise comparisons that include the tenth trajectory. We performed this analysis with  
4 the Python package *ProDy*.<sup>61</sup> Five snapshots of the tenth trajectory are shown in Figure  
5 5. The peptide starts to unfold within the first 6 ns and keeps the same partially unfolded  
6 structure after 40 ns, approximately. Thus, the peptide eventually attains this conformation,  
7 which seems stable, in most trajectories. The helical structure that remains is formed by  
8 six residues (Ala-Arg-Glu-Ala-Ala-Ala). The three consecutive alanines are localised at the  
9 centre of the peptide. Alanine has a propensity to form and stabilise  $\alpha$ -helix structure.<sup>62</sup>  
10 This property and central location likely explain the high stability of this part of the peptide  
11  $\alpha$ -helix conformation.  
12  
13  
14  
15  
16  
17  
18  
19  
20



23  
24  
25  
26  
27  
28  
29  
30  
31  
32 Figure 5: Typical trajectory from the 40 simulations.  
33  
34  
35

36 We also identified other trajectories where the unfolding process is in a more advanced  
37 state after 120 ns. One of these trajectories is shown in Figure 6. In this case, the peptide  
38 is completely unfolded after 80 ns. Although the structure of the  $\alpha$ -helix is modified after  
39 80 ps, it only starts to unfold after 200 ps. The conformation at 10 ns is similar to the  
40 stable conformation we identified in Figure 5. However, in this case, the peptide continues  
41 to unfold and the unfolding process is completed after 80 ns, approximately. The trajectory  
42 in Figure 6 is less frequent than the type shown in Figure 5. It occurs only once in the set of  
43 40 trajectories. We can gather more detailed information about the unfolding process using  
44 a Ramachandran plot representation (Figure 7) showing each residue from the trajectory  
45 presented in Figure 6.  
46  
47  
48  
49  
50  
51  
52  
53  
54

55 Figure 7 allows us to obtain information about the secondary structure of each residue.<sup>63</sup>  
56  
57  
58  
59  
60

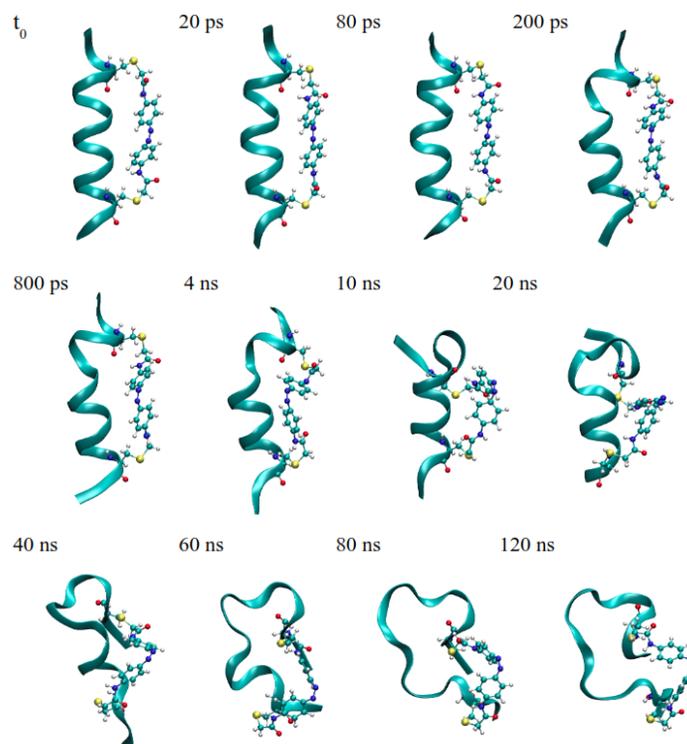


Figure 6: A trajectory where the peptide is completely unfolded by the end of the simulation.

The two most represented secondary structures are the  $\alpha$ -helix and the  $\beta$ -strand. They correspond to the cluster of points centered around  $\phi = -57^\circ$ ,  $\psi = -47^\circ$  and  $\phi = -80^\circ$ ,  $\psi = 150^\circ$ , respectively. As expected, all the residues form an  $\alpha$ -helix structure at the beginning of the simulation. The dihedral angles of the residues Arg2, Glu6, Ala8 and Ala15 are the first to change. They are followed by the residues Ala9, Ala10, Arg12 and Cys14, which switch from the helical region to the extended structure region of the Ramachandran plot. The dihedral angles of the residues Arg2, Cys3, Ala8 and Ala15 return approximately to their initial value at the end of the simulation. We also observe that the dihedral angles of the residues Arg2, Ala9, Ala10, and Arg12 transiently populate the upper-right quadrant of the Ramachandran plot, which corresponds to the region of the left-handed  $\alpha$ -helix structure.

In order to quantify the rate of the unfolding process we examine some representative quantitative progress coordinates, such as the distance between the two sulfur atoms and the  $O_i-H_{i+4}$  distances. The S-S distance is directly related to the structure of the azo-moiety.

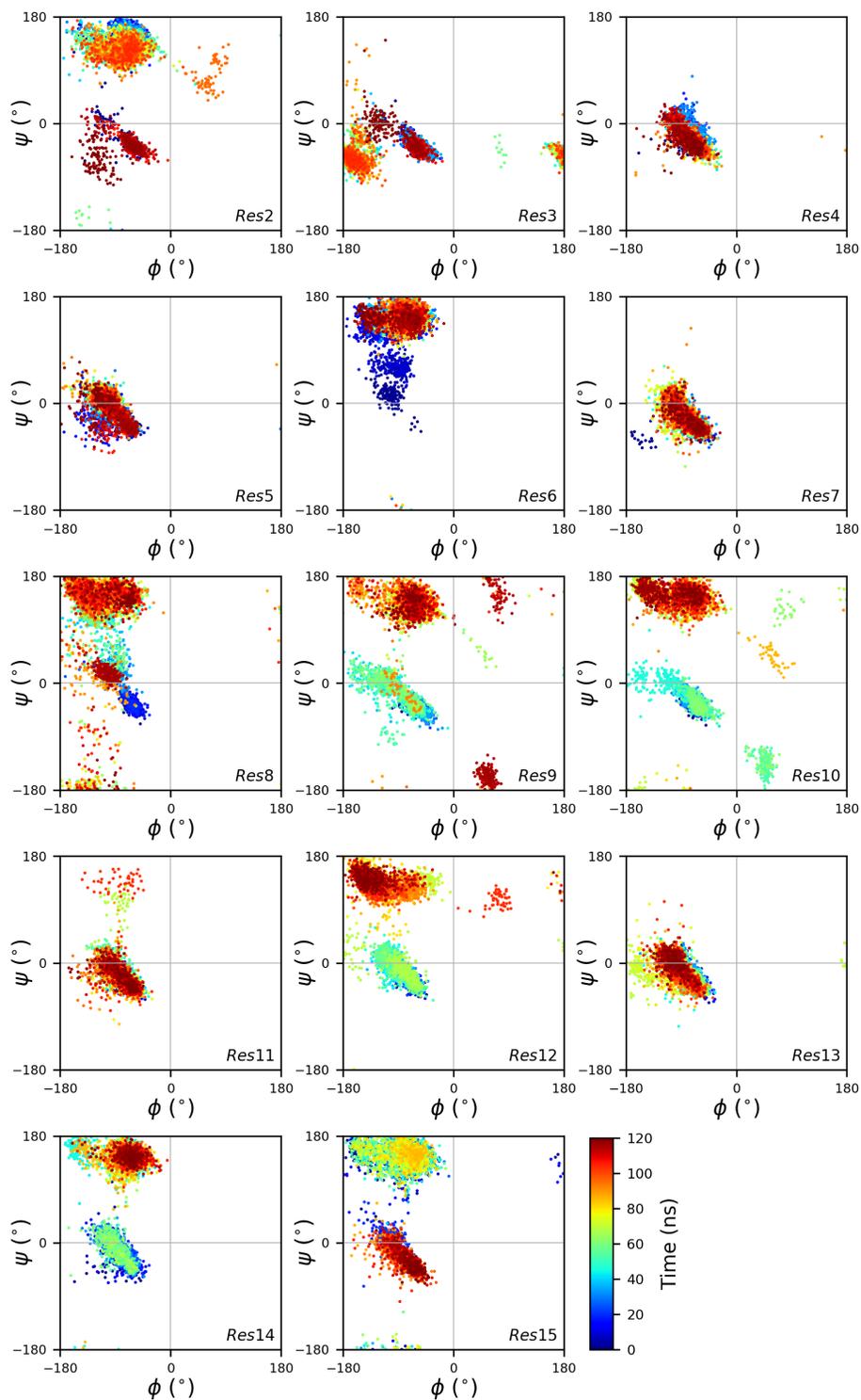


Figure 7: Ramachandran plots of the trajectory shown in Figure 6. The colour of the scatter plot corresponds to the time of the simulation.

We expect that, once isomerised, the azo-moiety tends to a lower energy conformation with this configuration. We also think that as long as the distance between the two sulfur atoms decreases, the azo-moiety induces a strain on the peptide. The hydrogen bond between  $O_i$  and  $H_{i+4}$  characterises the  $\alpha$ -helix structure and the distance between these two atoms allows one to assess the stability of the structure.<sup>64</sup> The averaged  $O_i$ - $H_{i+4}$  distance of the peptide residues will probably be partially correlated to the distance between the two sulfur atoms but we can expect that the conformation of the peptide continues to change even after the S-S distance converges. The S-S distance and the  $O_i$ - $H_{i+4}$  distances are presented in Figures 8 and 9, respectively.

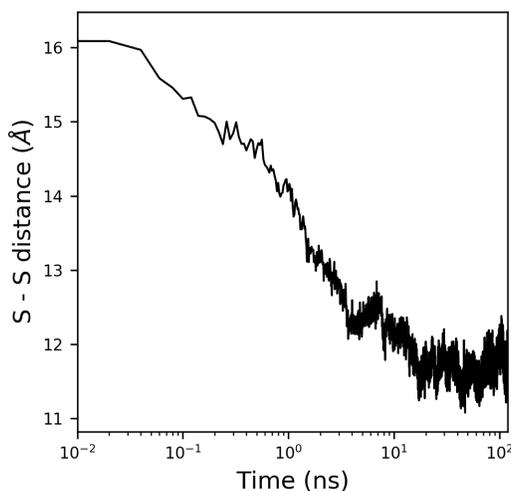


Figure 8: Distance between the two sulfur atoms of the cysteines over time (logarithmic scale).

In Figure 8, the S-S distance decreases from 16 Å to 12 Å within the first 10 ns of the simulation. Then it stabilises between 11 Å and 12 Å until the end of the simulation. We see in Figure 9a the heterogeneity in the speed of unfolding over the 40 trajectories. The trajectory in Figure 6 is trajectory number 26. Its mean  $O_i$ - $H_{i+4}$  distance (averaged over residues) is plotted against time in Figure 9b. The averaged  $O_i$ - $H_{i+4}$  distance from the 40 trajectories with the *cis* azo-moiety is shown in Figure 9c. The abscissa scale is logarithmic so the  $O_i$ - $H_{i+4}$  distance increases exponentially. The averaged  $O_i$ - $H_{i+4}$  distance increases

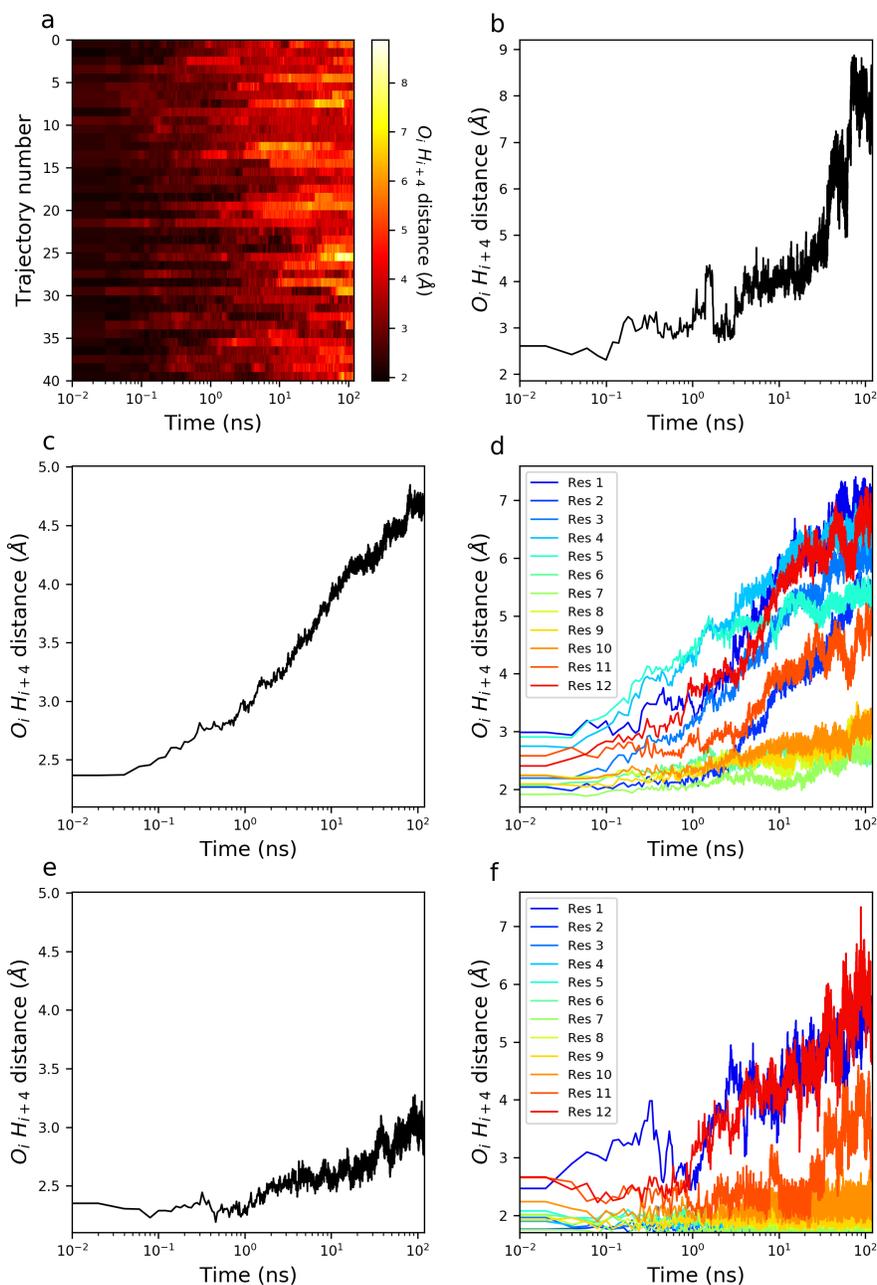


Figure 9: **a.** 2D representation of the averaged  $O_i-H_{i+4}$  distance for each trajectory over time. **b.** Averaged  $O_i-H_{i+4}$  distance of the trajectory shown in Figure 6 over time. **c.** Averaged  $O_i-H_{i+4}$  distance of the 40 trajectories with the *cis* azo-moiety over time. **d.** Averaged  $O_i-H_{i+4}$  distance of the 40 trajectories with the *cis* azo-moiety per residue over time. **e.** Averaged  $O_i-H_{i+4}$  distance of the 10 trajectories with the *trans* azo-moiety over time. **f.** Averaged  $O_i-H_{i+4}$  distance of the 10 trajectories with the *trans* azo-moiety per residue over time.

1  
2  
3 after 10 ns even though the S-S distance is constant. Thus, the peptide structure continues  
4 to change. In Figure 9b we see that the averaged  $O_i-H_{i+4}$  distance when the peptide is com-  
5 pletely unfolded is approximately 7.5 Å. Considering that Figure 9c shows that the averaged  
6  $O_i-H_{i+4}$  distance increases by approximately 1 Å every order of magnitude and reaches 4.5  
7 Å after 100 ns, it suggests that all the peptides should be completely unfolded after 100  
8  $\mu$ s. This extrapolated value does not correspond to values determined experimentally, which  
9 range between 100 ns and 1  $\mu$ s.<sup>40,41</sup> We will continue this comparison with experimental  
10 measurements in the next section. In Figure 9d, we plot the averaged  $O_i-H_{i+4}$  distance for  
11 the first 12 residues. This representation confirms that the  $\alpha$ -helix structure of the residues  
12 located at the centre of the peptide is more stable than at its termini. Earlier, we presented  
13 the hypothesis that the FK-11-X reaches a stable conformation after 40 ns approximately.  
14 We see in Figure 9d that the  $O_i-H_{i+4}$  distances of the residues Ala10 to Glu6 continue to  
15 increase even after 40 ns. It could mean that the stable conformation shown in Figure 5  
16 might be a local minimum and that the peptide will finally unfold completely on a longer  
17 time scale. The control simulation is shown in Figures 9e and 9f, which show the averaged  
18  $O_i-H_{i+4}$  distances and the averaged  $O_i-H_{i+4}$  distances, respectively, per residue from the ten  
19 trajectories with the *trans* azo-moiety. These two figures show that the termini of the  $\alpha$ -helix  
20 structure are not stable on the nanosecond time scale. However, we see a clear difference  
21 between the central residues with the *cis* and with the *trans* azo-moiety, which confirm the  
22 impact of the isomerisation of the azo-moiety on the peptide structure.  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

43 We performed an analysis based on CD spectra computed with the online version of  
44 Dichrocalc.<sup>23</sup> We calculated the CD spectra of the structures from MD simulations at 11  
45 different times (0, 0.02, 0.08, 0.2, 0.8, 2, 8, 20, 30, 60, and 120 ns). We selected these  
46 time points to sample homogeneously an exponential decay. The spectra of the trajectory  
47 illustrated in Figure 6 are shown in Figure 10. The CD spectrum at  $t_0$  is close to the spectral  
48 signature of an  $\alpha$ -helix with a positive peak at 190 nm and a negative peak between 200 and  
49 230 nm, centered at 208 nm. Whilst the well-known band at 222 nm appears as a shoulder  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

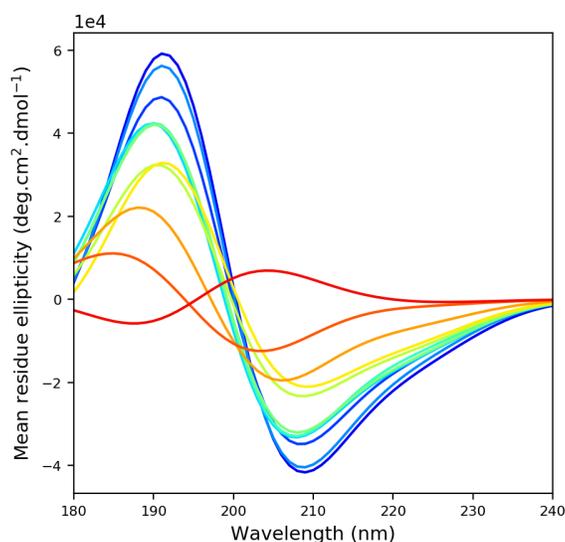


Figure 10: CD spectra calculated from the trajectory shown in Figure 6. The spectra belong to the structures extracted at 0 (blue), 0.02, 0.08, 0.2, 0.8, 2, 8, 20, 30, 60, and 120 ns (red).

rather than being more clearly distinct from the 208 nm band, the calculated intensity at 222 nm has been previously shown to be highly correlated with the experimental intensity at that wavelength.<sup>11</sup> The amplitude of the signal decreases progressively over the simulation. We also noticed that the spectrum of the last time frame, at 120 ns, is different. It exhibits a negative peak at 190 nm and a positive peak at 205 nm. This spectrum corresponds to the spectral signature of a random coil. Figure 6 shows that the  $\alpha$ -helix content of the peptide decreases over the simulation and it is completely unfolded after 80 ns. In the literature, the stable conformation of the *cis* FK-11-X peptide has been associated with the conformation at 120 ns in Figure 6.<sup>39</sup> The experimental synchrotron radiation CD (SRCD) spectra of *trans* and the *cis* configuration of FK-11-X have the same shape but with different amplitudes<sup>65</sup>(Figure 11b). Thus, the stable conformation of the *cis* FK-11-X is partially helical. If it were not, its CD spectra would correspond to a random coil based structure as the computed CD spectra at 120 ns (Figure 10). In addition, Auvray *et al.*<sup>65</sup> recently showed that the use of the natural polarisation of the synchrotron radiation may improve the temporal resolution of the technique to reach sub-nanosecond resolution. They assessed

the capacity of the instrument to record changes in CD of small molecules by monitoring the change in concentration of the enantiomers of the FK-11-X peptide. This novel SRCD setup would allow one to follow ultrafast dynamics occurring within a few tens of nanoseconds and thus it would also help to assess the reliability of computational models such as the one presented in this paper by providing transient SRCD spectra.

In order to analyse results that represent all the trajectories, we averaged the computed CD spectra from the 40 simulations to obtain averaged CD spectra that may be compared to the experimental SRCD spectra in solution. The computed CD spectra at  $t = 0$  ns and  $t = 120$  ns and the associated averaged spectra are plotted in Figure 11a. Approximately 80% of the trajectories populated an ensemble of related structured conformations and the others are between this set of structures and a completely unfolded ensemble. We see that the impact of the helical content on the spectrum is more important than the unfolded structure because the averaged spectrum at 120 ns still corresponds to an  $\alpha$ -helix. In addition, the

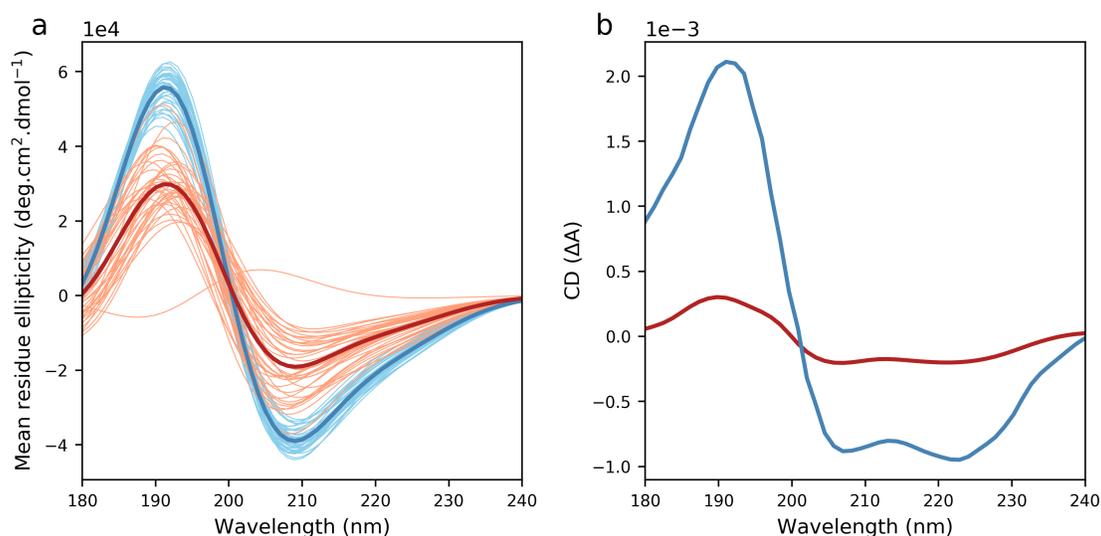


Figure 11: **a.** The computed CD spectra of the 40 trajectories from structures extracted at  $t = 0$  ns (blue) and  $t = 120$  ns (red). The thicker curves correspond to the average spectra. **b.** SRCD spectra of *cis* (red) and *trans* (blue) FK-11-X, as published in ref [63] under a Creative Commons CC BY license.

SRCD spectra plotted in Figure 11b show that the amplitude of the CD signal at 222 nm decreases by 80% when the configuration of the azo-moiety passes from *trans* to *cis*. The

1  
2  
3 same decrease in amplitude of the averaged CD spectrum computed with the structures at  
4  $t = 0$  ns (*trans* configuration) would result in a computed CD spectrum with an amplitude of  
5  $-3,600$  deg cm<sup>2</sup> dmol<sup>-1</sup> at 222 nm. The amplitude of the average of the CD spectra computed  
6 with the structures at  $t = 120$  ns (*cis* configuration) is  $-9,800$  deg cm<sup>2</sup> dmol<sup>-1</sup> at 222 nm.  
7 Thus, it decreases by 46% within the first 120 ns of simulation. The differences may come  
8 from either the sampling and could be addressed by increasing the number of trajectories or  
9 the duration of the simulation. We observed in Figure 9b that the averaged O<sub>*i*</sub>-H<sub>*i+4*</sub> distance  
10 continues to increase at the end of the simulation, so we can expect that the average helical  
11 content of the 40 trajectories tends to the value determined experimentally within the next  
12 tens of nanoseconds.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

23 In Figure 11a we see that in some trajectories, the peptide is in a more advanced unfolding  
24 state, in which the CD at 222 nm decreases to  $-5,000$  deg cm<sup>2</sup> dmol<sup>-1</sup>, approximately,  
25 which corresponds to a decrease of 70%. This estimation is close to the value determined  
26 with experimental results and thus the computed structure of the peptide at 120 ns in  
27 these trajectories may correspond to the most likely conformation of the FK-11-X in the *cis*  
28 configuration. The structures from three of these trajectories are shown in Figure 12. The  
29 helical content that remains in these structures is formed by the same residues as the helix  
30 shown in Figure 5, which include four alanines (Ala-Arg-Glu-Ala-Ala-Ala). Again, it shows  
31 the propensity of alanines to stabilise  $\alpha$ -helix structure.  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 Chen *et al.*<sup>40</sup> used data from time-resolved ORD experiments to estimate a time constant  
42 of  $55 \pm 6$  ns (single-exponential decay) for the unfolding process of the FK-11-X peptide.  
43 This estimation is based on data points measured from 50 ns to 10  $\mu$ s after the trigger of  
44 the isomerisation. The lack of information over the first 50 ns leads to a significant over-  
45 estimation of the value of the time constant. Based on this time-constant estimation, we can  
46 deduce that the dynamics should reach 95% of its total variation after  $165 \pm 6$  ns. So, even if  
47 this estimation may be too long, it shows that the unfolding process occurs within the first  
48  $165 \pm 6$  ns after the isomerisation of the azobenzene. In order to estimate the rate of the  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

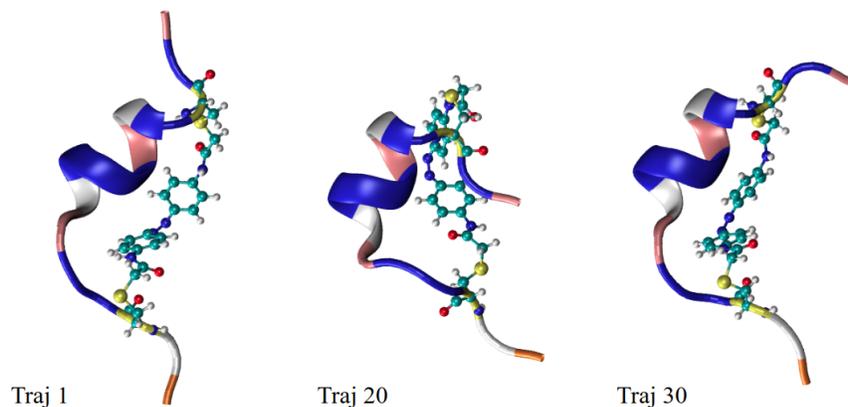


Figure 12: Structure of the FK-11-X peptide at 120 ns of the trajectories 1, 20 and 30. The colors of the peptide correspond to the amino acid type. Gln, Arg, Cys, Ala and Glu are in orange, white, yellow, blue and pink, respectively.

unfolding process via the change in CD, we averaged the computed CD spectra from the 40 simulations at 0, 0.02, 0.08, 0.2, 0.8, 2, 8, 20, 30, 60, and 120 ns. The averaged computed CD spectra and the trace of the CD signal at 222 nm are plotted in Figures 13a and 13b. The average spectra show that the amplitude of the CD signal decreases continuously over the simulation and homogeneously between 180 and 240 nm. The time-evolution of the intensity at 222 nm (Figure 13b) is sigmoidal, albeit not sharply so, reflecting a cooperative process, which evident in the correlated increase of several of the  $O_i-H_{i+4}$  distances seen in Figure 9d. In Figure 13b, the signal decreases rapidly over the first 20 ns and it starts to stabilise after 20 ns. So we can expect that the signal continues to decrease slightly over the next tens of nanoseconds. We performed a non-weighted least-squares fitting with a single-exponential decay model of the computed CD signal at 222 nm. The resulting curve is plotted in Figure 13b. This fitting allows us to estimate the time-constant of the simulated unfolding process at  $\tau=5.80\pm 0.03$  ns. Because we see that the computed CD signal seems to not reach a plateau at the end of the simulation, this may be an under-estimate and could be refined with longer simulations or perhaps with simulations using a polarisable force field. The final slower decrease in amplitude expected after 120 ns should make the loss of amplitude at 222 nm approach 80%, the value that we deduced from experimental data. These results show

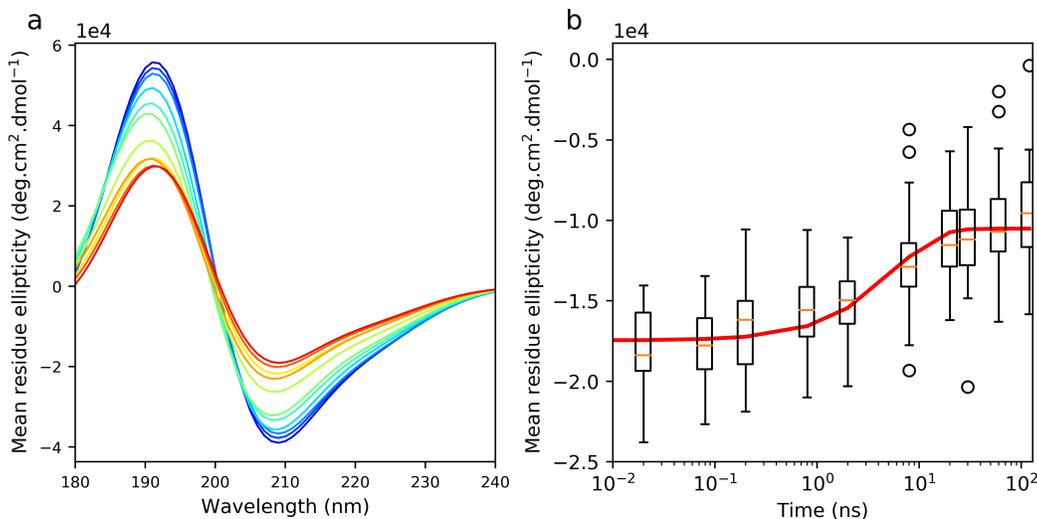


Figure 13: **a.** Averaged CD spectra of the 40 trajectories from structures extracted at 0 (blue), 0.02, 0.08, 0.2, 0.8, 2, 8, 20, 30, 60, and 120 ns (red). **b.** Trace of the computed CD signal at 222 nm. The boxes extend from the first to the third quartile values of the data, with a horizontal red line at the median. The whiskers attached to the boxes extend to the extrema of the data. If the whiskers are 50% longer than the difference between the first and the third quartile values, then the outer data points are represented as circles. The red curve is a non-weighted least-squares fitting with a single-exponential decay model with a time constant of  $\tau=5.80\pm 0.03$  ns.

that the rate of the simulated unfolding process is consistent with that provided by time-resolved ORD experiments. Comparison with the time-resolved ORD experiment of Chen *et al.*<sup>40</sup> is more relevant than comparison with the time-resolved IR data from Bredenbeck *et al.*,<sup>41</sup> because the *trans* to *cis* isomerization is followed by equilibration on the *cis* energy landscape, whereas the latter investigated dynamics on the *trans* landscape.

Khuc *et al.*<sup>6</sup> performed quantitative time-resolved CD measurement with a time step of 12 ns at 225 nm. They successfully estimated the variation of the helical fraction of a poly(glutamic acid) peptide over a T-jump experiment using a nanosecond Nd:YAG laser to heat up the solvent and a sub-picosecond source based on a 1 kHz amplified Titanium-Sapphire system for the CD measurement.<sup>66</sup> Dartigalongue *et al.*<sup>66</sup> used this light source to perform sub-picosecond measurements to assess the change in CD over the photodissociation of carbonmonoxy-myoglobin. Thus, time-resolved pump-probe CD measurements with either sub-picosecond or nanosecond temporal resolution on the FK-11-X peptide using

1  
2  
3 this approach would bring further information which would complement the model that we  
4 propose in this paper.  
5  
6  
7

## 8 9 **Conclusion**

10  
11 We have investigated the unfolding of the FK-11-X peptide based on a computational ap-  
12 proach. We optimised a force field for the azo-moiety to ensure that this part of the FK-11-X  
13 peptide is parameterised well and in an internally consistent fashion. Our parameters should  
14 find utility in future simulation studies of other azo-peptides. Using this force field, we com-  
15 puted 40 MD trajectories of 120 ns simulating the unfolding process of the FK-11-X peptide.  
16 Approximately 80% of the trajectories presented a partially unfolded peptide with a similar  
17 conformation at the end of the simulation. The remaining 20% were in a more advanced  
18 unfolded state and in one trajectory the peptide was completely unfolded. We analysed  
19 the dynamics of the process looking at the distance between the two sulfur atoms and the  
20  $O_i-H_{i+4}$  backbone hydrogen bond distances. The S-S distance decreases progressively during  
21 the first 10 ns and then is stable until the end of the simulation. The  $O_i-H_{i+4}$  distances  
22 and the structure of the peptide at 120 ns indicated that the three alanines located between  
23 the two cysteines significantly stabilise the  $\alpha$ -helix. We also see that the termini of the  
24 FK-11-X helix are quite labile on the nanosecond timescale, even in the *trans* configuration.  
25 The amino acids at the termini unfold progressively over the first 100 ns. An analysis of  
26 CD spectra computed from snapshots from the MD simulations allowed us to compare the  
27 dynamics of the simulated unfolding process and the experimental data, with the caveat  
28 that an improved model of the coupling of the  $\pi\pi^*$  and  $n\pi^*$  transitions might resolve the  
29 calculated band at 222 nm better. The calculated CD signal at 222 nm reaches a plateau at  
30 the end of the simulation. The  $55 \pm 6$  ns time constant proposed by Chen *et al.* allows us  
31 to estimate that 95% of the reaction would occur within  $165 \pm 6$  ns. This estimate matches  
32 well with the rate of the simulated dynamics. We also used the variation of the CD signal  
33 at 222 nm from experimental data and from the calculated spectra to demonstrate that the  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 most likely stable conformation of the *cis* FK-11-X is not the fully unfolded conformation  
4  
5 but a conformation close to the structure shown in Figure 12.  
6  
7

## 8 9 **Supporting Information Available**

10  
11 The supporting information includes an assessment of the equilibration process and compares  
12  
13 the force field parameters that we developed for the azobenzene with those proposed by  
14  
15 Böckmann *et al.*<sup>51</sup>  
16  
17

## 18 19 **Acknowledgments**

20  
21 We thank the University of Nottingham and SOLEIL for a Ph.D. studentship for F.A. We  
22  
23 are grateful for access to the University of Nottingham High Performance Computer.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- (1) Brown, J. E.; Klee, W. A. Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **1971**, *10*, 470–476.
- (2) Scholtz, J. M.; Baldwin, R. L. The mechanism of alpha-helix formation by peptides. *Ann. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–118.
- (3) Andrew, C. D.; Bhattacharjee, S.; Kokkoni, N.; Hirst, J. D.; Jones, G. R.; Doig, A. J. Stabilizing interactions between aromatic and basic side chains in alpha-helical peptides and proteins. Tyrosine effects on helix circular dichroism. *J. Am. Chem. Soc.* **2002**, *124*, 12706–12714.
- (4) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* **1996**, *35*, 691–697.
- (5) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. Alpha-helix peptide folding and unfolding activation barriers: A nanosecond UV resonance Raman study. *J. Am. Chem. Soc.* **1999**, *121*, 8074–8086.
- (6) Khuc, M. T.; Mendonca, L.; Sharma, S.; Solinas, X.; Volk, M.; Hache, F. Measurement of circular dichroism dynamics in a nanosecond temperature-jump experiment. *Rev. Sci. Instrum.* **2011**, *82*, 054302.
- (7) Chen, Y. H.; Yang, J. T.; Martinez, H. M. Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* **1972**, *11*, 4120–4131.
- (8) Woody, R. W. Circular dichroism. *Meth Enzymol.* **1995**, *246*, 34–71.
- (9) Christov, C.; Tielens, F.; Mirazchiiski, M. Modeling study of the influences of the

- 1  
2  
3 aromatic transitions and the local environment on the far-UV rotational strengths in  
4 TEM-1 beta-lactamase. *J. Mol. Model.* **2006**, *12*, 411–416.  
5  
6  
7
- 8 (10) Dang, Z.; Hirst, J. D. Short hydrogen bonds, circular dichroism and over-estimates of  
9 peptide helicity. *Angew. Chemie Intl. Ed.* **2001**, *113*, 3731–3733.  
10  
11  
12
- 13 (11) Besley, N. A.; Hirst, J. D. Theoretical studies toward quantitative protein circular  
14 dichroism calculations. *J. Am. Chem. Soc.* **1999**, *121*, 9636–9644.  
15  
16  
17
- 18 (12) Shea, J. E.; Brooks III, C. L. From folding theories to folding proteins: A review and  
19 assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys.*  
20 *Chem.* **2001**, *52*, 499–535.  
21  
22  
23
- 24 (13) Georgoulia, P. S.; Glykos, N. M. Molecular simulation of peptides coming of age: Ac-  
25 curate prediction of folding, dynamics and structures. *Arch. Biochem. Biophys.* **2019**,  
26 *664*, 76–88.  
27  
28  
29
- 30  
31 (14) Rogers, D. M.; Jasim, S. B.; Dyer, N. T.; Auvray, F.; Réfrégiers, M.; Hirst, J. D.  
32 Electronic circular dichroism of proteins. *Chem* **2019**, *5*, 2751–2774.  
33  
34  
35
- 36 (15) Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y. H.; Goto, Y.; Réfrégiers, M.; Kardos, J.  
37 Accurate secondary structure prediction and fold recognition for circular dichroism  
38 spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E3095–E3103.  
39  
40  
41  
42
- 43 (16) Mavridis, L.; Janes, R. W. PDB2CD: A web-based application for the generation of  
44 circular dichroism spectra from protein atomic coordinates. *Bioinformatics* **2017**, *33*,  
45 56–63.  
46  
47  
48
- 49 (17) Nagy, G.; Igaev, M.; Jones, N. C.; Hoffmann, S. V.; Grubmuller, H. SESCO: Predicting  
50 circular dichroism spectra from protein molecular structures. *J. Chem. Theor. Comput.*  
51 **2019**, *15*, 5087–5102.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (18) Kaminský, J.; Kubelka, J.; Bour, P. Theoretical modeling of peptide *alpha*-helical cir-  
4 cular dichroism in aqueous solution. *J. Phys. Chem. A* **2011**, *115*, 1734–1742.  
5  
6  
7  
8 (19) Seibert, J.; Bannwarth, C.; Grimme, S. Biomolecular structure information from high-  
9 speed quantum mechanical electronic spectra calculation. *J. Am. Chem. Soc.* **2017**,  
10 *139*, 11682–11685.  
11  
12  
13 (20) Woody, R. W. Improved Calculation of the  $n\pi^*$  Rotational Strength in Polypeptides.  
14 *J. Chem. Phys.* **1968**, *49*, 4797–4806.  
15  
16  
17 (21) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. Rotatory properties of molecules con-  
18 taining two peptide groups: theory. *J. Phys. Chem.* **1969**, *73*, 228–243.  
19  
20  
21 (22) Ianeselli, A.; Orioli, S.; Spagnolli, G.; Faccioli, P.; Cupellini, L.; Jurinovich, S.; Men-  
22 nucci, B. Atomic detail of protein folding revealed by an ab initio reappraisal of circular  
23 dichroism. *J. Am. Chem. Soc.* **2018**, *140*, 3674–3682.  
24  
25  
26 (23) Bulheller, B. M.; Hirst, J. D. DichroCalc—circular and linear dichroism online. *Bioin-*  
27 *formatics* **2009**, *25*, 539–540.  
28  
29  
30 (24) Hildebrand, N.; Michaelis, M.; Wurzler, N.; Li, Z.; Hirst, J. D.; Micsonai, A.; Kar-  
31 dos, J.; Gil-Ley, A.; Bussi, G.; Koppen, S., et al. Atomistic details of chymotrypsin  
32 conformational changes upon adsorption on silica. *ACS Biomater. Sci. Eng.* **2018**, *4*,  
33 4036–4050.  
34  
35  
36 (25) Michaelis, M.; Hildebrand, N.; Meißner, R. H.; Wurzler, N.; Li, Z.; Hirst, J. D.; Mic-  
37 sonai, A.; Kardos, J.; Delle Piane, M.; Colombi Ciacchi, L. Impact of the conforma-  
38 tional variability of oligopeptides on the computational prediction of their CD spectra.  
39 *J. Phys. Chem. B* **2019**, *123*, 6694–6704.  
40  
41  
42  
43  
44 (26) First, J. T.; Webb, L. J. Agreement between experimental and simulated circular  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- dichroic spectra of a positively charged peptide in aqueous solution and on self-assembled monolayers. *J. Phys. Chem. B* **2019**, *123*, 4512–4526.
- (27) Gattuso, H.; Garcia-Iriepa, C.; Sampedro, D.; Monari, A.; Marazzi, M. Simulating the electronic circular dichroism spectra of photoreversible peptide conformations. *J. Chem. Theor. Comput.* **2017**, *13*, 3290–3296.
- (28) Szymański, W.; Beierle, J. M.; Kistemaker, H. A. V.; Velema, W. A.; Feringa, B. L. Reversible photocontrol of biological systems by the incorporation of molecular photo-switches. *Chem. Rev.* **2013**, *113*, 6114–6178.
- (29) Woolley, G. A. Photocontrolling peptide  $\alpha$  helices. *Acc. Chem. Res.* **2005**, *38*, 486–493.
- (30) Beharry, A. A.; Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chem. Soc. Rev.* **2011**, *40*, 4422–4437.
- (31) Zhao, Y.; Ikeda, T. *Smart light-responsive materials: azobenzene-containing polymers and liquid crystals.*; John Wiley & Sons, 2009.
- (32) Nenov, A.; Borrego-Varillas, R.; Oriana, A.; Ganzer, L.; Segatta, F.; Conti, I.; Segarra-Marti, J.; Omachi, J.; Dapor, M.; Taioli, S., et al. UV-light-induced vibrational coherences: the key to understand Kasha rule violation in trans-azobenzene. *J. Phys. Chem. Lett.* **2018**, *9*, 1534–1541.
- (33) Nägele, T.; Hoche, R.; Zinth, W.; Wachtveitl, J. Femtosecond photoisomerization of cis-azobenzene. *Chem. Phys. Lett.* **1997**, *272*, 489–495.
- (34) Fujino, T.; Tahara, T. Picosecond time-resolved Raman study of trans-azobenzene. *J. Phys. Chem. A* **2000**, *104*, 4203–4210.
- (35) Fujino, T.; Arzhantsev, S. Y.; Tahara, T. Femtosecond time-resolved fluorescence study of photoisomerization of trans-azobenzene. *J. Phys. Chem. A* **2001**, *105*, 8123–8129.

- 1  
2  
3 (36) Renner, C.; Moroder, L. Azobenzene as conformational switch in model peptides. *Chem-*  
4 *BioChem* **2006**, *7*, 868–678.  
5  
6  
7  
8 (37) Spörlein, S.; Carstens, H.; Satzger, H.; Renner, C.; Behrendt, R.; Moroder, L.; Ta-  
9 van, P.; Zinth, W.; Wachtveit, J. Ultrafast spectroscopy reveals subnanosecond pep-  
10 tide conformational dynamics and validates molecular dynamics simulation. *Proc. Natl.*  
11 *Acad. Sci. U.S.A.* **2002**, *99*, 7998–8002.  
12  
13  
14  
15  
16 (38) Buchli, B.; A., W. S.; Walser, R.; Donten, M. L.; Pfister, R.; Blöchliger, N.; Steiner, S.;  
17 Cafilisch, A.; Zerbe, O.; Hamm, P. Kinetic response of a photoperturbed allosteric  
18 protein. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 11725–11730.  
19  
20  
21  
22  
23 (39) Flint, D. G.; Kumita, J. R.; Smart, O. S.; Woolley, G. A. Using an azobenzene cross-  
24 linker to either increase or decrease peptide helix content upon trans-to-cis photoiso-  
25 merization. *Chem. Biol.* **2002**, *9*, 391–397.  
26  
27  
28  
29  
30 (40) Chen, E.; Kumita, J. R.; Woolley, G. A.; Kliger, D. S. The kinetics of helix unfolding of  
31 an azobenzene cross-linked peptide probed by nanosecond time-resolved optical rotatory  
32 dispersion. *J. Am. Chem. Soc.* **2003**, *125*, 12443–12449.  
33  
34  
35  
36  
37 (41) Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P.  $\alpha$ -Helix formation  
38 in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved  
39 IR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2379–2384.  
40  
41  
42  
43  
44 (42) Xia, S. H.; Cui, G.; Fang, W. H.; Thiel, W. How photoisomerization drives peptide  
45 folding and unfolding: Insights from QM/MM and MM dynamics simulations. *Angew.*  
46 *Chem. Int. Ed.* **2016**, *128*, 2107–2112.  
47  
48  
49  
50  
51 (43) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchi-  
52 son, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis  
53 platform. *J. Cheminformatics* **2012**, *4*, 17.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (44) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol.*  
4 *Graph.* **1996**, *14*, 33–38.  
5  
6  
7  
8 (45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.  
9 Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*  
10 **1983**, *79*, 926–935.  
11  
12  
13  
14 (46) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.;  
15 Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with  
16 NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.  
17  
18  
19  
20  
21 (47) MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.;  
22 Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S., et al. All-atom empirical potential  
23 for molecular modeling and dynamics studies of proteins. *J. Phys. Chem.* **1998**, *102*,  
24 3586–3616.  
25  
26  
27  
28  
29  
30 (48) MacKerell Jr, A. D.; Feig, M.; Brooks, C. L. Improved treatment of the protein back-  
31 bone in empirical force fields. *J. Am. Chem. Soc.* **2003**, *126*, 698–699.  
32  
33  
34  
35 (49) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D.  
36 Optimization of the additive CHARMM all-atom protein force field targeting improved  
37 sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi^1$  and  $\chi^2$  dihedral angles. *J. Chem.*  
38 *Theory Comput.* **2012**, *8*, 3257–3273.  
39  
40  
41  
42  
43  
44 (50) Carstens, H. Konformationsdynamik lichtschtbarer peptide: Molekulardynamiksimu-  
45 lationen und datengetriebene modellbildung [Conformational dynamics of photoswitch-  
46 able peptides: Molecular dynamics and data-driven model building]. PhD thesis (Lud-  
47 wig Maximilians Universität München, Munich). Available at: [http://edoc.ub.uni-](http://edoc.ub.uni-muenchen.de/2268)  
48 [muenchen.de/2268](http://edoc.ub.uni-muenchen.de/2268). **2004**,  
49  
50  
51  
52  
53  
54 (51) Böckmann, M.; Peter, C.; Site, L. D.; Doltsinis, N. L.; Kremer, K.; Marx, D. Atomistic  
55  
56  
57  
58  
59  
60

- 1  
2  
3 force field for azobenzene compounds adapted for QM/MM simulations with applica-  
4 tions to liquids and liquid crystals. *J. Chem. Theory Comput.* **2007**, *3*, 1789–1802.  
5  
6  
7
- 8 (52) Vanommeslaeghe, K.; MacKerell Jr, A. D. Automation of the CHARMM General Force  
9 Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52*,  
10 3144–3154.  
11  
12  
13
- 14 (53) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Dar-  
15 ian, E.; Guvench, O.; Lopes, P.; Vorobyov, I., et al. CHARMM general force field: A  
16 force field for drug-like molecules compatible with the CHARMM all-atom additive  
17 biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.  
18  
19  
20  
21  
22
- 23 (54) Ihalainen, J. A.; Bredenbeck, J.; Pfister, R.; Helbing, J.; Chi, L.; van Stokkum, I. H.;  
24 Woolley, G. A.; Hamm, P. Folding and unfolding of a photoswitchable peptide from  
25 picoseconds to microseconds. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5383–5388.  
26  
27  
28  
29
- 30 (55) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.;  
31 Lange, A. W.; Behn, A.; Deng, J.; Feng, X., et al. Advances in molecular quantum  
32 chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–  
33 215.  
34  
35  
36  
37  
38
- 39 (56) Feller, S. E.; Pastor, R. W.; Rojnuckarin, A.; Bogusz, S.; Brooks, B. R. Effect of  
40 electrostatic force truncation on interfacial and transport properties of water. *J. Phys.*  
41 *Chem.* **1996**, *100*, 17011–17020.  
42  
43  
44  
45
- 46 (57) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE  
47 algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962.  
48  
49  
50
- 51 (58) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian  
52 equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J.*  
53 *Comput. Phys.* **1977**, *23*, 327–341.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (59) Paoli, B.; Seeber, M.; Backus, E. H.; Ihalainen, J. A.; Hamm, P.; Caffisch, A. Bulky  
4 side chains and non-native salt bridges slow down the folding of a cross-linked helical  
5 peptide: a combined molecular dynamics and time-resolved infrared spectroscopy study.  
6 *J. Phys. Chem. B* **2009**, *113*, 4435–4442.  
7  
8  
9  
10  
11 (60) Amadei, A.; Ceruso, M. A.; Di Nola, A. On the convergence of the conformational  
12 coordinates basis set obtained by the essential dynamics analysis of proteins' molecular  
13 dynamics simulations. *Proteins Str. Func. Bioinf.* **1999**, *36*, 419–424.  
14  
15  
16  
17 (61) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: protein dynamics inferred from theory  
18 and experiments. *Bioinformatics* **2011**, *27*, 1575–1577.  
19  
20  
21  
22 (62) Spek, E. J.; Olson, C. A.; Shi, Z.; Kallenbach, N. R. Alanine is an intrinsic  $\alpha$ -helix  
23 stabilizing amino acid. *J. Am. Chem. Soc.* **1999**, *121*, 5571–5572.  
24  
25  
26  
27 (63) Hollingsworth, S. A.; Karplus, P. A. A fresh look at the Ramachandran plot and the  
28 occurrence of standard structures in proteins. *Biomol. Concepts* **2010**, *1*, 271–283.  
29  
30  
31  
32 (64) Hirst, J. D.; Brooks III, C. L. Molecular dynamics simulations of isolated helices of  
33 myoglobin. *Biochemistry* **1995**, *34*, 7614–7621.  
34  
35  
36  
37 (65) Auvray, F.; Dennetiere, D.; Giuliani, A.; Jamme, F.; Wien, F.; Nay, B.; Zirah, S.;  
38 Polack, F.; Menneglier, C.; Lagarde, B.; Hirst, J. D.; Réfrégiers, M. Time resolved  
39 transient circular dichroism spectroscopy using synchrotron natural polarization. *Struct.*  
40 *Dynam.* **2019**, *6*, 054307.  
41  
42  
43  
44 (66) Dartigalongue, T.; Niezborala, C.; Hache, F. Subpicosecond UV spectroscopy of  
45 carbonmonoxy-myoglobin: absorption and circular dichroism studies. *Phys. Chem.*  
46 *Chem. Phys.* **2007**, *9*, 1611–1615.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

# Supporting Information

## Unfolding Dynamics of a Photo-switchable Helical Peptide

Francois Auvray and Jonathan D. Hirst

### 1. Equilibration RMSD

The averaged root-mean-squared deviation (RMSD) of the peptide backbone is plotted as a function of time in Figure S1.

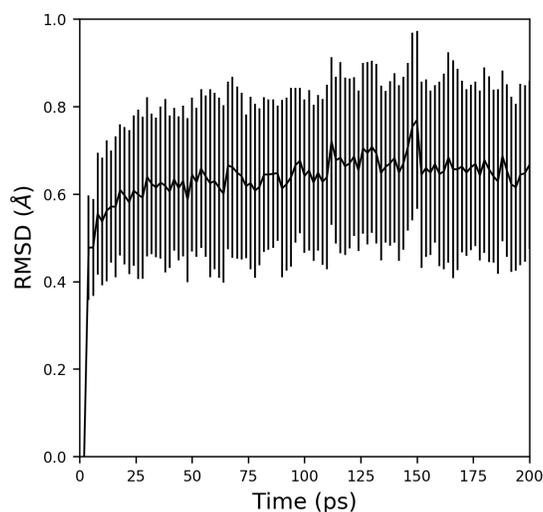


Figure S1: RMSD variation over time. The error bars width corresponds to one standard deviation. The RMSD calculation is only based on the atoms of the peptide backbone.

The reference of our RMSD calculation was the initial conformation of the equilibration simulation. The RMSD increases over the first 25 ps of the simulation and then stabilises around 0.65 Å, showing that the backbone structures reach stable conformations and that the systems are equilibrated.

## 2. Force field parameters

The labeled structures of the parameterised molecules are shown in Figure S2. The force field parameters proposed by Böckmann *et al.*<sup>51</sup> for the azobenzene (Figure S2a) are shown in Table S1. The optimised parameters presented in this paper for the structure shown in Figure S2a and S2b are in Table S2 and S3, respectively.

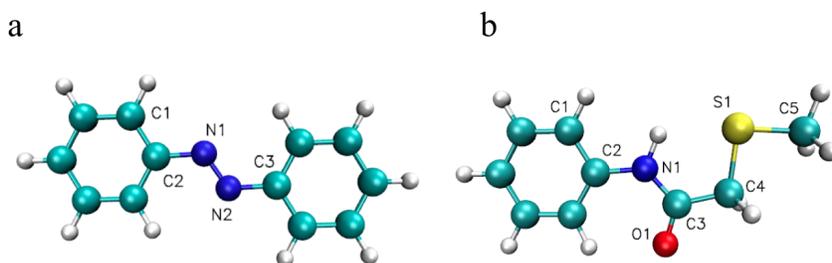


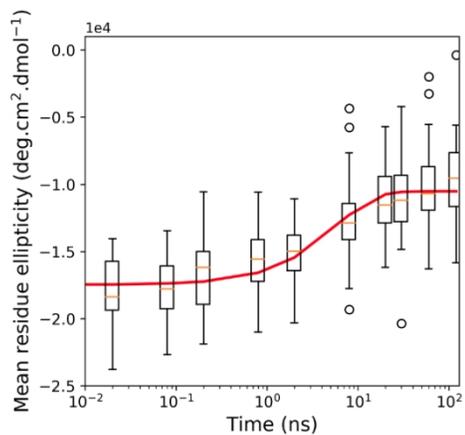
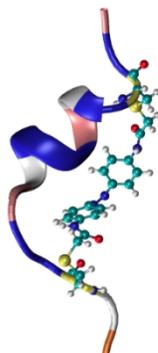
Figure S2: **a.** Azobenzene **b.** Structures used for the force field parameter optimisation of the linker between the azobenzene and the peptide.

Table S1: Optimised force field parameters of the azobenzene (Figure S2a) proposed by Böckmann *et al.*<sup>51</sup>

| <b>Dihedral angle</b>                                       | $k$ (kcal/mol) | $n$ | $\delta$ ( $^\circ$ ) | <b>Angle</b>                                 | $k$ (kcal/mol) | $\theta_0$ ( $^\circ$ ) | <b>Bond</b>                   | $k$ (kcal/mol) | $b_0$ ( $\text{\AA}$ ) |
|---|----------------|-----|-----------------------|--|----------------|-------------------------|-------------------------------|----------------|------------------------|
| C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 1.43           | 2   | 180                   | C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> | 134            | 120.0                   | C <sub>2</sub> N <sub>1</sub> | 172            | 1.432                  |
| C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> C <sub>3</sub> | 16.70          | 2   | 180                   | C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 155            | 116.5                   | N <sub>1</sub> N <sub>2</sub> | 335            | 1.262                  |

Table S2: Optimised force field parameters of the azobenzene (Figure S2a).

| <b>Dihedral angle</b>                                       | $k$ (kcal/mol) | $n$ | $\delta$ ( $^\circ$ ) | <b>Angle</b>                                 | $k$ (kcal/mol) | $\theta_0$ ( $^\circ$ ) | <b>Bond</b>                   | $k$ (kcal/mol) | $b_0$ ( $\text{\AA}$ ) |
|---|----------------|-----|-----------------------|--|----------------|-------------------------|-------------------------------|----------------|------------------------|
| C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 2.5            | 2   | 180                   | C <sub>1</sub> C <sub>2</sub> N <sub>1</sub> | 50             | 116.5                   | C <sub>2</sub> N <sub>1</sub> | 300            | 1.415                  |
| C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> C <sub>3</sub> | 23.0           | 2   | 180                   | C <sub>2</sub> N <sub>1</sub> N <sub>2</sub> | 75             | 106.0                   | N <sub>1</sub> N <sub>2</sub> | 500            | 1.295                  |



TOC graphic

338x190mm (96 x 96 DPI)