ELECTRONIC SUPPLEMENTARY INFORMATION:

The energy landscape of Aβ_{42}: a funnel to disorder for the monomer becomes a folding funnel for self-assembly

Moritz Schäffler,[a,b] David J. Wales,[c] and Birgit Strodel[a,b,*]

1 Methods

1.1 Simulation Details

In this study, molecular dynamics (MD) simulations were conducted for the Aβ_{1-42} monomer and dimer. In both systems, Aβ_{1-42} was modeled with neutral histidine and no terminal capping groups, resulting in an overall peptide charge of 3−. The simulations were performed using the GROMACS simulation package.[1] As the monomer and dimer simulations where originally performed as part of different studies, some of the simulation settings differ slightly, but the most relevant settings are identical, in particular the force field parameters and ion concentration are the same. We are confident that the slight differences in the setup do not affect our key conclusions.

The CHARMM36m force field[2] was employed in all simulations. Previous research has revealed that the CHARMM36m force field provides an accurate representation of monomeric Aβ[3] and is well suited for simulating amyloid aggregation phenomena.[4]

All system preparations followed the same protocol: the peptide(s) was(were) positioned within the simulation box with a minimum distance of 1.2 nm between any peptide atom and the simulation box faces or edges. The box was subsequently filled with TIP3P water molecules,[5] along with Na^+ and Cl^- ions to achieve system neutralization and a physiological salt concentration of 150 mM. After system equilibration, each system was simulated under NpT conditions at 1 bar, using the Parrinello-Rahman pressure coupling scheme.[6] The dimer simulations were conducted at 298 K using the Nosé-Hoover thermostat,[7,8] while the monomer system was maintained at 300 K using a velocity rescaling thermostat.[9] In all simulations, periodic boundary conditions in all dimensions were applied, with the particle-mesh Ewald method[10] employed for the calculation of electrostatic interactions. The calculation of van der Waals and Coulomb interactions in real space were performed with a cutoff at 1.2 nm.

The total simulation time for the Aβ_{1-42} monomer accumulated to 6 μs. For the dimer, three simulations were conducted, each initiated from two extended Aβ_{1-42} monomers separated by at least 2 nm. Due to the initial peptide extended conformations, a relatively large simulation box was necessary to prevent self-interaction across periodic boundary conditions, leading to an extensive system predominantly composed of water. To conserve computational resources, after an initial 2 μs simulation per dimer simulation, the peptide structures of the last MD frame were extracted and resolvated in a smaller simulation box, as at this point the two peptides had already formed a dimer. Subsequently, after a brief NpT equilibration, another 6 μs per dimer system were collected, resulting in a cumulative simulation time of 3×8 μs for the Aβ_{1-42} dimer.

All MD simulations were run on the supercomputer JURECA.[11]

1.2 Distance of Reciprocal Interatomic Distances Metric

In order to partition the sampled conformational space over the course of the MD simulation into discrete microstates, we used the distribution of reciprocal interatomic distance (DRID) metric[12] for subsequent structural clustering. Given the size of the Aβ_{1-42} monomer, structure-based clustering in Cartesian coordinates becomes a challenging task. One therefore usually relies on some form of dimensionality reduction, which should preserve as much of the kinetics and structural features as possible. It has been shown that the DRID metric is a good candidate to meet both of these criteria.[12,13] A key feature of the DRID metric is the use of the multiplicative inverse (reciprocal) distances, which highlights the difference in short distances, while not neglecting changes in large distances completely.

To apply the DRID metric, two essential atom sets are defined: a set of m centroids $\mathscr{C}$ representing key structural elements, and a set of N reference atoms $\mathscr{A}$ (excluding atoms that are covalently bound to the centroid). For a given structure, the distribution of reciprocal interatomic distances for each centroid i and the first three moments of that distribution (μ_i, ν_i, ξ_i) are calculated, resulting in a $3m$ dimensional vector for each structure (i.e., each frame of the MD trajectory). The moments are defined as follows,

$$\mu_i = \frac{1}{N-1-nb_i} \sum_i^N \frac{1}{d_{ij}}, \tag{1}$$

$$\nu_i = \left[\frac{1}{N-1-nb_i} \sum_i^N \frac{1}{(d_{ij}-\mu_i)^2} \right]^{1/2}, \tag{2}$$

[a] Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[b] Institute of Biological Information Processing, Structural Biochemistry (IBI-7), Forschungszentrum Jülich, 52428 Jülich, Germany
[c] Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, U.K.
* Correspondence: b.strodel@fz-juelich.de

$$\xi_i = \left[\frac{1}{N-1-nb_i} \sum_i^N \frac{1}{\left(d_{ij}-\mu_i\right)^3} \right]^{1/3}, \quad (3)$$

where d_{ij} denotes the distance of atom $a_j \in \mathscr{A}$ to centroid $c_i \in \mathscr{C}$, nb_i is the number of covalent bonds of a centroid, and N the number of atoms in $\mathscr{A}$. The distance metric s_{jk} between a pair of conformations j and k in DRID space is defined as

$$s_{jk} = \frac{1}{3m} \sum_i^m \left[(\mu_i^j - \mu_i^k)^2 + (\nu_i^j - \nu_i^k)^2 + (\xi_i^j - \xi_i^k)^2 \right]^{1/2}. \quad (4)$$

To group the structures into states, we performed regular space clustering in DRID space, as implemented in the PyEMMA python package,[14] using the s_{jk} distance metric.

For studying the conformational space of the Aβ_{1-42} peptide, we chose as centroids the C_α atoms of structurally important residues, namely D1, F19, D23, K28, L34 and A42, resulting in an 18-dimensional DRID space. We chose D23 and K28, because they have been identified in previous studies to form a salt bridge in the β-hairpin state.[15,16] The residues F19 and L34 are members of the hydrophobic core regions and have been shown by NMR spectroscopy to form contacts in Aβ oligomers,[17] while the termini D1 an A42 where included to capture the overall compactness of the peptide. For clustering, we chose a cutoff of $s_{jk}^c = 0.02$ nm^{-1}, which resulted in 447 states for the monomer and 511 states for the dimer.

To evaluate the robustness of the FES derived from projecting structures into the DRID space, we computed the FES for the dimer system using various DRID metrics. Specifically, we selected five times 6 random residues of the peptide chain as centroids for defining the DRID metric, and constructed five FESs. All FESs exhibited a consistent single funnel structure, similar to the FES presented in this manuscript (Fig. 2). To assess the predictability of our results, we calculated the overlap between the ensemble of states corresponding to the global minimum of the FES presented in Fig. 2 and those of the randomly generated FESs, yielding an average overlap of ∼45%. When the three most prominent states next to the global minimum but within the same basin are further included, there is an overlap of at least 77% between the current and the random FES. Moreover, the intrinsically disordered state consistently appeared as a high-energy excited state in all FESs. These findings underscore the robustness of the DRID metric, showing its reliability relatively independent of centroid selection.

1.3 Free Energy Calculation

The free energy surface (FES) of a protein determines its structural and dynamical properties and is therefore of great interest if one wants to understand the protein function. Here, we calculate the free energies associated with the states determined by structural clustering in DRID space, treating each state as a minimum in the FES.[18] The free energy of the minima F_i are calculated via their occupation probability p_i,

$$F_i = -k_B T \log(p_i), \quad (5)$$

where k_B is the Boltzmann constant and T the temperature of the system. The rate matrix $\mathbf{R}$, representing the state-to-state rates r_{jk} between minima j and k observed in the MD simulations, is used to derive the transition state free energies F_{jk} via the Eyring–Polanyi formulation. The rate matrix was derived from the transition matrix of the MD trajectory in DRID space, which represents the corresponding right stochastic matrix. Assuming Markovian dynamics, these free energies can be calculated as follows:

$$F_{jk} = F_k - k_B T \log(k_{jk}) + k_B T \log\left(\frac{k_B T}{h}\right), \quad (6)$$

where h is the Planck constant. In a perfectly converged system, the transition state free energy for both interconversion rates between minima j and k should be the same, i.e. $F_{jk} = F_{kj}$. However, for a finite MD trajectory this equality is rarely achieved. To minimize the error we average over both rates,

$$F_{jk}^{\mathrm{ts}} = \frac{F_{jk} + F_{kj}}{2}, \quad (7)$$

giving us an estimate for the transition state free energy F_{jk}^{ts} between minima j and k.

It is important to note that it is the rates that correspond to observable quantities, and the free energy barriers simply provide a convenient way to visualise the corresponding landscape. Equation (6) translates the rates into barriers on a log scale. Multiplying all the rates by a constant factor would preserve detailed balance and shift the connections in the disconnectivity graph uniformly, without affecting the organisation of the landscape. The relative free energies of the minima in the graph and the barriers between them reproduce the equilibrium distribution and the rates by construction. Hence we obtain insight into the organisation of the landscape in a representation that will faithfully reproduce the stationary distribution and the rates.

To assess the assumption of Markovian dynamics, we conducted a Chapman-Kolmogorov test on the kinetic network of the dimer system, following the standard PyEMMA protocol[19] with a discrete timestep of $\tau = 20$ ps. The number of metastable states was set to five, representing the number of states on the pathway from the disordered state D to the global minimum B. Figure S5 illustrates the comparison between the estimated and predicted transition probabilities between metastable states for lagtimes up to 100 ps. The strong agreement between estimated and predicted probabilities supports the Markovian nature of the dynamics. Notably, the states involved in the pathway do not correspond to the metastable states identified by PCCA++ clustering[20]:

D : state 1

I_1 : state 3

I_2 : state 5

I_3 : state 5

B : state 5

However, this observation aligns with our findings, as states I_2, I_3,

and B are structurally similar, closely located within the same energy funnel, and exhibit relatively fast transitions, as confirmed by FPT analysis. Additionally, the highest transition probabilities are associated with the transitions between states $1 \rightarrow 3$ and $3 \rightarrow 5$, reproducing the fastest pathway as determined by FPT analysis. The estimated and predicted transition probabilities as determined from a Chapman-Kolmogorov test for the kinetic network of the monomer system is presented in Fig. S6.

1.4 First Passage Times

While the FES governs the structural and dynamical properties of a molecule, in experiment often the relaxation times associated with a given process are measured. Thus, studying the timescales associated with transitions between minima on the FES can bridge the gap between simulation and experiment, and also shed light on the processes associated with those timescales. Often quantified by the mean first passage time (MFPT), interconversion rates between minima offer insights into the average time taken for a system to transition between reactant and product states. It has been shown that studying the first passage time (FPT) distribution of a transition can reveal a interesting additional information.[21] Specifically, such analysis provides direct access to the organizational structure of the underlying energy landscape and facilitates the identification of distinct signatures linked to relaxation to different funnels in the FES.

For a given transition $\mathsf{A} \leftarrow \mathsf{B}$ from reactant state A to product state B the first passage time probability distribution $p(t)$ can be obtained by treating the product state as absorbing. Then, the master equation for the occupation probabilities $P_\alpha(t)$ for the set of intermediate states I and reactant states $\mathsf{I} \cup \mathsf{B}$ is

$$\begin{bmatrix} \dot{\mathbf{P}}_\mathsf{I}(t) \\ \dot{\mathbf{P}}_\mathsf{B}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{K}_\mathsf{II} - \mathbf{D}_\mathsf{I} & \mathbf{K}_\mathsf{IB} \\ \mathbf{K}_\mathsf{BI} & \mathbf{K}_\mathsf{BB} - \mathbf{D}_\mathsf{B} \end{bmatrix} \begin{bmatrix} \mathbf{P}_\mathsf{I}(t) \\ \mathbf{P}_\mathsf{B}(t) \end{bmatrix} = \mathbf{M}\mathbf{P}_{\mathsf{I}\cup\mathsf{B}}(t), \quad (8)$$

where $\mathbf{K}_{\mathsf{XY}}$ is the rate matrix of transitions between connected states and $\mathbf{D}_\mathsf{X}$ a diagonal matrix containing the escape rates of each state in X, i.e. $\lceil \mathbf{D}_\mathsf{X} \rceil_{ii} = \sum_j \mathbf{K}_{ji}$. Applying eigenvector decomposition to the formal solution of eq. 8 produces an analytic solution for the first passage time distribution

$$p(t) = \sum_l \nu_l e^{-\nu_l t} A_l, \quad (9)$$

here $-\nu_l$ are the eigenvalues of $\mathbf{M}$ and A_l are amplitudes, which depend on the eigenvectors of $\mathbf{M}$. The FPT can be represented by the probability distribution $\mathscr{P}(y)$ for $y = \log(t)$

$$\mathscr{P}(y) = \sum_l \nu_l e^{y - \nu_l \exp(y)} A_l, \quad (10)$$

which gives distinct peaks for different features and relaxation time scales in the FPT distribution.

2 Pathway Analysis

2.1 Extraction of Pathways

From the database of minima and transition states constituting the FES of the dimer, we extracted the shortest pathway connecting states D and state B using the PATHSAMPLE program.[22] This pathway describes the fastest transition from a disordered structure (state D) to a β-hairpin structure (state B). The transition occurs through three intermediate states (I_1, I_2, I_3), characterized by the establishment of a salt bridge between residues D23 and K28, followed by the formation of hydrophobic contacts and subsequent rearrangement of these contacts into a β-sheet. However, due to the dimensionality reduction from application of the DRID metric before clustering, the resulting minima represent an ensemble of configurations. While these structures are similar according to the DRID metric, this procedure leads to an ensemble of slightly different pathways. To identify potential pathways from this ensemble of candidates for further analysis, we have employed linear interpolation between the configurations to construct trajectories for the transition:

$$\mathrm{D} \rightarrow \mathrm{I}_1 \rightarrow \mathrm{I}_2 \rightarrow \mathrm{I}_3 \rightarrow \mathrm{B}.$$

Given our focus on studying the pathway associated with the fastest transition rates, we only considered trajectories that do not involve chain crossing in the linear interpolation between states, which is an artefact.. From this procedure, we inferred three pathways (Fig. S2), which exhibit slight differences in their final configuration but all begin with the formation of a salt bridge between D23 and K28, proceed with hydrophobic contact formation, and ultimately reorient into a β-sheet configuration.

The predominant pathway 1 is discussed in the main text. Here, we briefly discuss the other two pathways from D to B. Pathway 2 is different from pathway 1 primarily in two aspects. Firstly, the closing scissor motion from an extended to a closed structure is not fully accomplished via the transition to the first intermediate state I_1, but also involves the second intermediate state I_2. Nonetheless, it follows the same hierarchy of events, where first the salt bridge and then the hydrophobic contacts are formed. The second deviation from pathway 1 is the orientation of the C-terminus, which almost wraps around the rest of the chain. Examination of the full dimer structure of this member of the global minimum ensemble reveals a very stable complex, which is the predominant structure in one of the three MD trajectories. In contrast, pathway 3 exhibits β-hairpin formation only in the early stages of the transition, via the first intermediate state I_1. Subsequently, the hydrophobic contacts reorganise to form a short parallel β-sheet in I_1, resulting in a wide loop in the peptide central region. While this pathway differs significantly from pathways 1 and 2, the final structure constitutes only a very small fraction of the equilibrium population of the global minimum ($< 1\%$).

2.2 First Passage Times

Besides calculating the FPT probability distribution for the overall transition, we also calculated the FPTs between the intermediate states corresponding to the fastest pathway between D and B. Fig. S4 shows the resulting FPT probability distributions for each intermediate transition. In each case, the FPT probability distribution indicates the presence of both fast and slow relaxation modes to the target state, leading to an overall slow MFPT. The two peaks correspond to two competing events: relaxation to the

global minimum and relaxation to the absorbing target state of that transition. The first, faster peak is associated with direct relaxation to the absorbing target state, while the slower peak corresponds to relaxation to the global minimum followed by eventual propagation to the absorbing target state. This is evident in the FPT probability distribution, where slow relaxation to the global minimum is generally more likely, except for the final transition $I_3 \rightarrow B$, where fast/direct relaxation to the global minimum is more probable compared to relaxation via intermediates. The last intermediate exhibits significant structural similarity to the global minimum, which is further underscored by the close proximity of I_3 and B in the FES funnel. Consequently, the slow peak corresponding to relaxation via intermediates in the side funnel acts only as a minor kinetic trap. This latter process is not prominently featured in the FPT distribution for the $D \rightarrow B$ transition, as it is not accessed significantly for the chosen starting point and thus has a negligible impact on the overall relaxation pathway.

The FPTs associated with the fastest mode according to the peak position of the fast relaxation mode are as follows:

$$\tau^{\mathrm{dim}}_{\mathrm{D}\rightarrow\mathrm{I}_1} \sim 780 \text{ ps}$$

$$\tau^{\mathrm{dim}}_{\mathrm{I}_1\rightarrow\mathrm{I}_2} \sim 190 \text{ ps}$$

$$\tau^{\mathrm{dim}}_{\mathrm{I}_2\rightarrow\mathrm{I}_3} \sim 90 \text{ ps}$$

$$\tau^{\mathrm{dim}}_{\mathrm{I}_3\rightarrow\mathrm{B}} \sim 110 \text{ ps}$$

Since each FPT probability distribution was calculated separately, the respective target state was treated as an absorbing state. As a result, the individual estimates of the FPT associated with the fastest transitions do not add up to the estimate of the overall transition time between states D and B of $\tau^{\mathrm{dim}}_{\mathrm{D}\rightarrow\mathrm{B}} \sim 4$ ns, which include all possible recrossing event. The individual FPT distributions provide insights into the relative timescales of the intermediate transitions as well as the underlying structure of the FES.

2.3 Cooperative Folding

We demonstrated how the FES of the Aβ_{1-42} peptide is significantly altered in the presence of another Aβ_{1-42} peptide by employing the same DRID metric to calculate the FES. However, this approach only implicitly reveals the impact of their interaction. To more thoroughly analyse the cooperative effects driving the Aβ_{1-42} peptide towards a more folded, β-sheet-rich state rather than a disordered state, we studied the intra- and interpeptide contacts along the fastest pathway from state D to B.

Figure S3 displays the intra- and interpeptide contact maps for the ensemble of states states belonging to D and B as well as for the three intermediates along that pathway. In the disordered state D, the two peptides show no interaction, as indicated by the absence of contacts between the two peptides. The intrapeptide contact map, on the other hand, reveals a slight tendency towards forming a hairpin structure, suggested by the faint contact trace perpendicular to the diagonal. This mirrors the behaviour of two Aβ_{1-42} monomers, as shown by the FES of the Aβ_{1-42} monomer in Fig. 1, and confirms that no stable β-hairpin is formed without an interaction partner. In the first intermediate state I_1, the two Aβ_{1-42} peptides begin to form a complex, with one monomer contacting the hydrophobic C-terminal region of the other. In parallel, contacts begin to form within the peptide, in particular a salt bridge forms between residues D23 and K28, and the hydrophobic core region around F19 interacts with the hydrophobic C-terminal region around L34. At this point, the orientation of the interacting peptide segments within the peptide is mainly parallel and not antiparallel, as required for a hairpin structure. In the second intermediate state I_2, the interpeptide contact map reveals strong interactions between the hydrophobic regions of the two Aβ_{1-42} peptides, while the intramolecular D23–K28 salt bridge is broken to allow the hydrophobic intrapeptide contacts to reorganise. The binding of the hydrophobic regions between the peptides can be considered a hydrophobic platform essential for forming the proper intrapeptide contacts necessary for the β-hairpin structure. In the third intermediate state I_3, the D23–K28 salt bridge reforms, and the β-hairpin structure becomes established, as indicated by a strong contact trace perpendicular to the diagonal. Additionally, the dimer shows a high propensity for forming antiparallel contacts between the hydrophobic core regions or hydrophobic C-termini of the two peptides. Finally, the transition to the global minimum state B is characterised by further stabilisation of the β-hairpin and a stronger tendency for antiparallel alignment in the intermolecular hydrophobic contacts.

In summary, this juxtaposition of intra- and interpeptide interactions during the $D \rightarrow B$ transition reveals that their formation occurs in a cooperative manner. The Aβ_{1-42} peptide gains the ability to fold into a stable β-hairpin by binding to the hydrophobic region offered by the other Aβ_{1-42} peptide. Therefore, the increase in hydrophobicity in the environment appears to be crucial for folding during self-assembly.

Notes and references

1 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19–25.
2 J. Huang, S. Rauscher, G. Nawrocki, R. Ting, M. Feig, B. de Groot, H. Grubmüller and A. MacKerell, *Nature Meth.*, 2017, **14**, 71–73.
3 A. Paul, S. Samantray, M. Anteghini, M. Khaled and B. Strodel, *Chem. Sci.*, 2021, **12**, 6652–6669.
4 S. Samantray, F. Yin, B. Kav and B. Strodel, *J. Chem. Inf. Model.*, 2021, **60**, 6462 – 6475.
5 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
6 M. Parrinello and A. Rahman, *Mol. Phys.*, 1981, **52**, 7182–7190.
7 S. Nosé, *Mol. Phys.*, 1984, **52**, 255–268.
8 W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695–1697.
9 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
10 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993.
11 D. Krause and P. Thörnig, *JLSRF*, 2018, **4**, A132.
12 T. Zhou and A. Caflisch, *J. Chem. Theory Comput.*, 2012, **8**, 2930–2937.
13 D. Chakraborty, J. E. Straub and D. Thirumalai, *Sci. Adv.*, 2023, **9**, eadd6921.
14 M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz and F. Noé, *J. Chem. Theory Comput.*, 2015, **11**, 5525–5542.
15 G. Reddy, J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 11948–11953.
16 M. Schäffler, S. Samantray and B. Strodel, *Int. J. Mol. Sci.*, 2023, **24**, 11238.
17 J. Jeon, W.-M. Yau and R. Tycko, *Nature Commun.*, 2023, **14**, 2964.
18 S. V. Krivov and M. Karplus, *J. Chem. Phys.*, 2002, **117**, 10894–10903.
19 M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz and F. Noé, *J. Chem. Theory Comput.*, 2015, **11**, 5525–5542.
20 S. Röblitz and M. Weber, *Adv. Data Anal. Classif.*, 2013, **7**, 147–179.
21 D. J. Wales, *J. Phys. Chem. Lett.*, 2022, **13**, 6349–6358.
22 PATHSAMPLE*: A program for generating connected stationary point databases and extracting global kinetics*, http://www-wales.ch.cam.ac.uk/software.html.

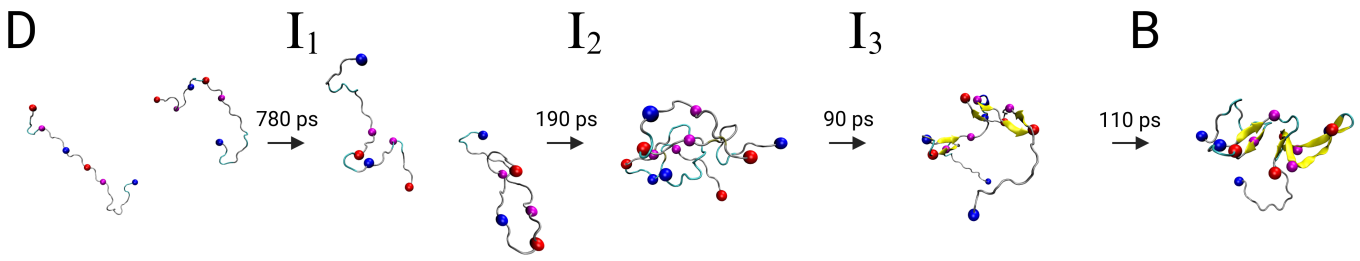

Fig. S1 Fastest pathway from intrinsically disordered state (D) to the global minimum (B) in the dimer free energy surface. Different to Fig. 3, here both peptides in each of the states are shown. The spheres represent the centroids used in the DRID metric: blue spheres for positively charged N-terminus and K28, red spheres for the negatively charged C-terminus and D23, and magenta spheres for the hydrophobic F19 and L34.

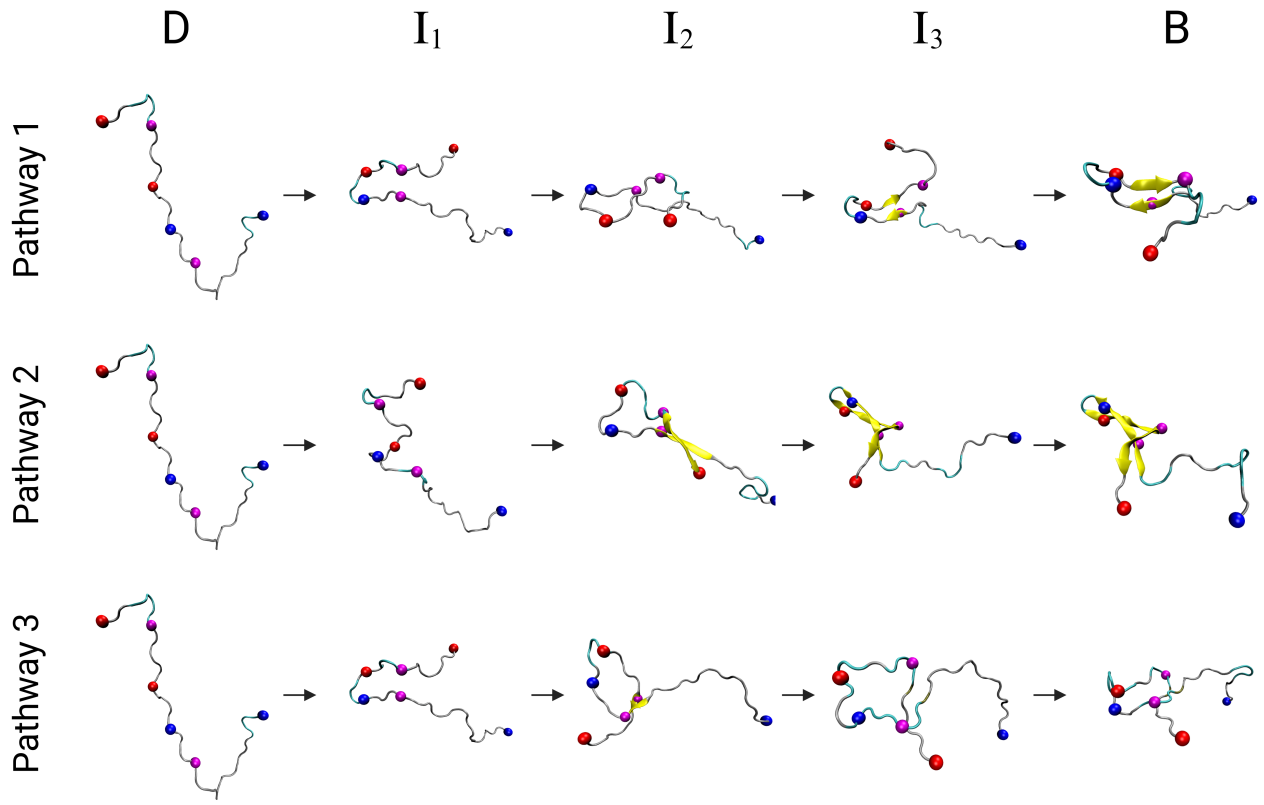

Fig. S2 The three most dominant pathways from the intrinsically disordered state (D) to the global minimum (B) in the free energy surface of the dimer. Only the peptide of the dimer for which the D $\rightarrow$ B transition was analysed is shown.

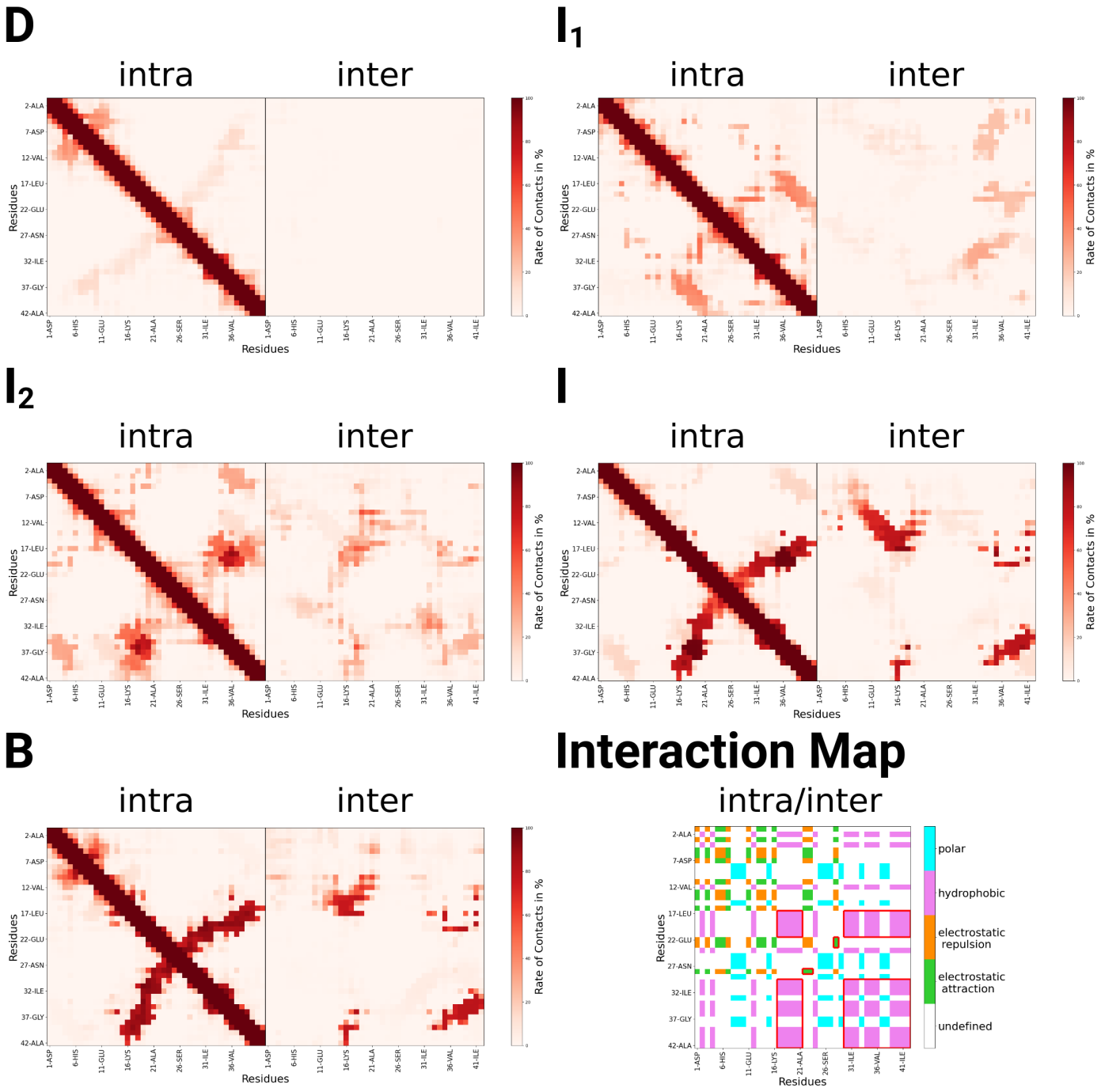

Fig. S3 Intra- and interpeptide contact maps along the D → B transition for Aβ_{1-42} dimerisation. The contact maps were calculated for the ensemble of conformations belonging to the intrinsically disordered state (D), the global minimum (B), and the intermediate states of the fastest D→B pathway (I_1,I_2,I_3). Two residues were considered to be in contact if any pair of atoms of the two residues were within a distance of 6 Å. The lower right corner displays the physicochemical nature of all possible residue–residue interactions that could occur within an Aβ_{1-42} peptide or between Aβ_{1-42} peptides: electrostatic repulsion (orange) or attraction (green), polar (blue) and hydrophobic interactions (magenta). For direct comparison with the other panels of this figure, the regions of relevant hydrophobic interactions as well as the D23–K28 salt bridge are highlighted by red boxes.

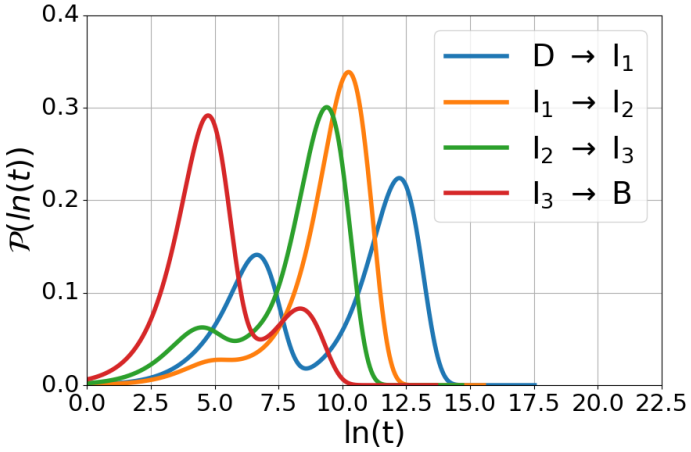

Fig. S4 Probability distribution of the first passage time t, $\mathscr{P}(\ln t)$, for each intermediate transition of the pathway $D \rightarrow I_1 \rightarrow I_2 \rightarrow I_3 \rightarrow B$. The FPTs are calculated for each transition separately, treating the target state as an absorbing state.

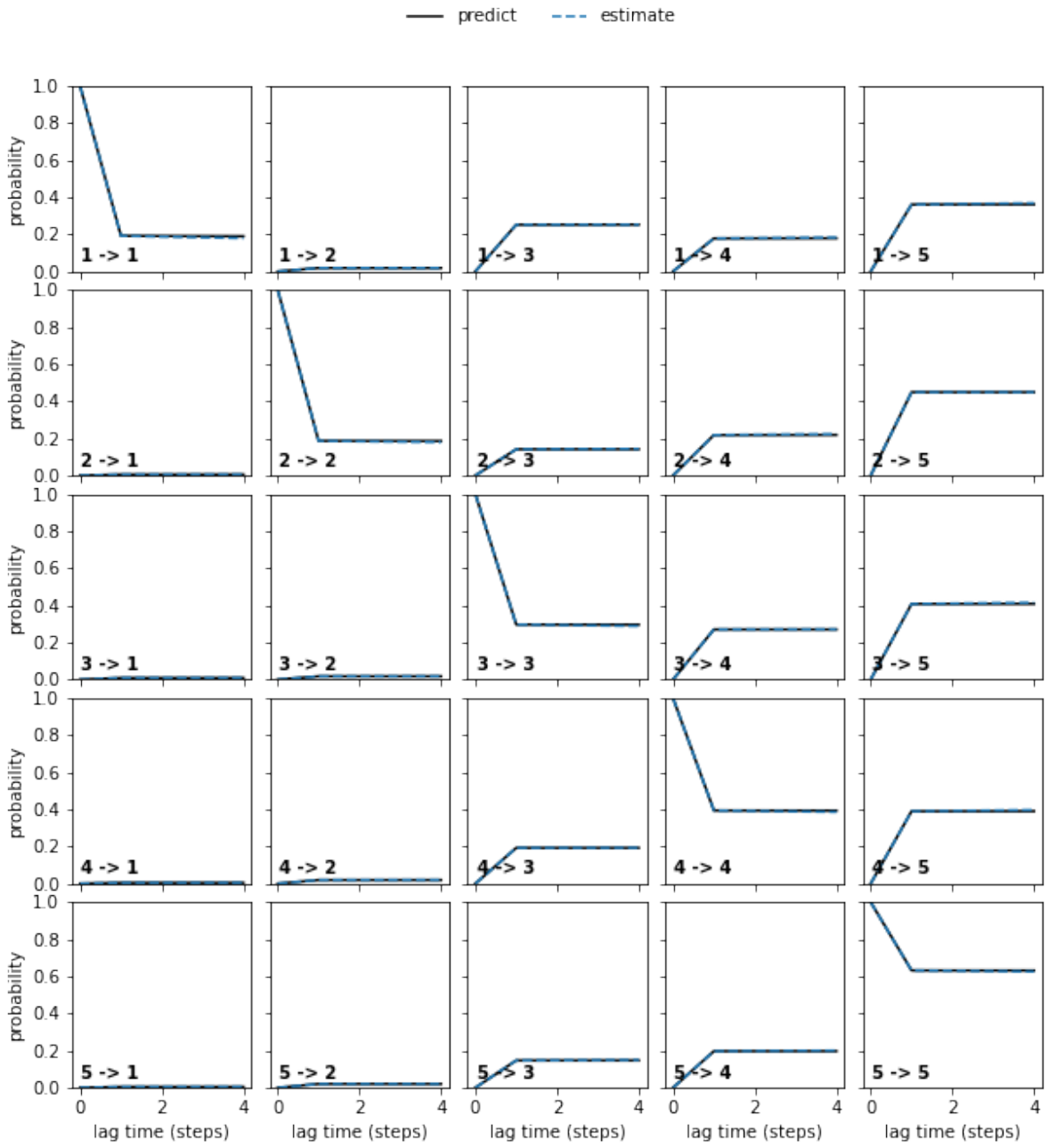

Fig. S5 Chapman-Kolmogorov test for the kinetic network of the dimer system, following a standard PyEMMA protocol with a discrete timestep of τ=20 ps. The number of metastable states as determined by PCCA++ was set to 5.

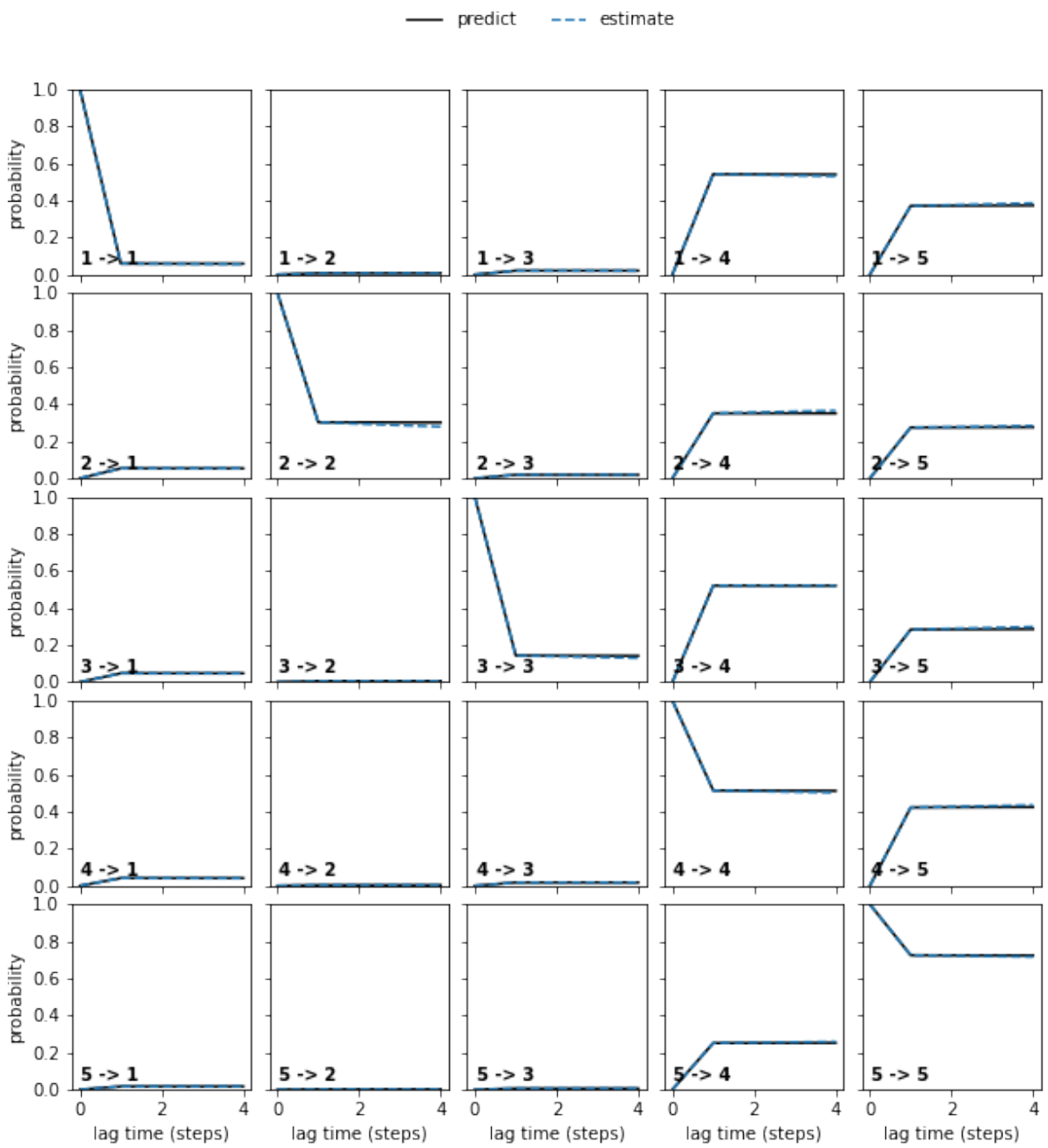

Fig. S6 Chapman-Kolmogorov test for the kinetic network of the monomer system, following a standard PyEMMA protocol with a discrete timestep of τ=20 ps. The number of metastable states as determined by PCCA++ was set to 5

.