International Journal of
Molecular Sciences

MDPI

Article

Energy Landscapes and Heat Capacity Signatures for Monomers and Dimers of Amyloid-Forming Hexapeptides

Nicy and David J. Wales *

Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK
* Correspondence: dw34@cam.ac.uk

Abstract: Amyloid formation is a hallmark of various neurodegenerative disorders. In this contribution, energy landscapes are explored for various hexapeptides that are known to form amyloids. Heat capacity (C_V) analysis at low temperature for these hexapeptides reveals that the low energy structures contributing to the first heat capacity feature above a threshold temperature exhibit a variety of backbone conformations for amyloid-forming monomers. The corresponding control sequences do not exhibit such structural polymorphism, as diagnosed via end-to-end distance and a dihedral angle defined for the monomer. A similar heat capacity analysis for dimer conformations obtained using basin-hopping global optimisation shows clear features in end-to-end distance versus dihedral correlation plots, where amyloid-forming sequences exhibit a preference for larger end-to-end distances and larger positive dihedrals. These results hold true for sequences taken from tau, amylin, insulin A chain, a de novo designed peptide, and various control sequences. While there is a little overall correlation between the aggregation propensity and the temperature at which the low-temperature C_V feature occurs, further analysis suggests that the amyloid-forming sequences exhibit the key C_V feature at a lower temperature compared to control sequences derived from the same protein.

Keywords: energy landscapes; amyloid formation; hexapeptides; heat capacity

check for updates

Citation: Nicy; Wales, D.J. Energy Landscapes and Heat Capacity Signatures for Monomers and Dimers of Amyloid-Forming Hexapeptides. *Int. J. Mol. Sci.* **2023**, *24*, 10613. https://doi.org/10.3390/ijms241310613

Academic Editor: Ludmilla A. Morozova-Roche

Received: 18 May 2023
Revised: 16 June 2023
Accepted: 20 June 2023
Published: 25 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Amyloid fibrils are associated with serious disorders, such as Alzheimer's, Parkinson's, type II diabetes, and dialysis-related amyloidosis. These amyloids are formed from specific proteins, and the oligomers are toxic [1]. Amyloid fibrils have a cross-beta structure that is formed by H-bonding between the NH and CO groups of the main chain of partially folded or misfolded proteins [2,3]. Since it is generally possible for proteins to establish such interactions between main chain atoms, amyloid formation has been suggested to be a generic property [4]. However, amino acid sequences also play an important role, and amyloid formation propensity may be a sequence-specific property [5,6]. Both amyloid-forming proteins and short segments of these proteins can form steric zippers [7–10]. All the peptides that form fibrils also form stable dimers [11]. The aim of the present work is to investigate whether the propensity for amyloid formation is encoded in the energy landscape of monomers and dimers, and to provide a thermodynamic diagnostic involving the heat capacity (C_V), which would complement models of amyloid formation based on the interactions within side chains.

The side chains of amino acid residues play an important role in amyloid formation [12–14], as they participate in aromatic–aromatic [15], electrostatic [16], and van der Waals interactions [17,18]. The hydrophobicity, secondary structure propensity, overall charge [19], exposed surface, dipole moment, and cooperativity in peptides correlate with amyloid formation ability [20]. The deposition of peptides on amyloid templates is stereospecific and it involves interactions between peptide backbone and/or side chains [21]. Several predictor algorithms have been developed to identify amyloidogenic

proteins based on the insights obtained from amyloid structure and properties of amino acid residues [22–25]. Different conformations of amyloid proteins can also occur at different temperatures and in different regions of the brain [26]. These alternative conformations result in amyloid polymorphism [27,28].

However, the thermodynamic driving force for amyloid formation is still not well understood. Amyloid formation is studied mainly from the kinetic perspective [29,30]. Several attempts have been made to study the heat capacity of proteins [31–33] using isothermal titration calorimetry and differential scanning calorimetry, and investigate the thermodynamics of amyloid formation [34–36]. Our recent heat capacity calculations have provided some insight into how the presence of tyrosine and arginine increases the phase separation propensity of proteins [37]. This approach is also useful to rationalise the context-dependent properties of amino acid residues in different sequences. Here, we explore the energy landscapes and calculate the heat capacity for monomers (Tables 1 and 2) and dimers (Table 1) of hexapeptides that are experimentally known to form amyloids, along with mutations that result in loss in amyloid formation ability (Figures 1 and 2). We find that the low-energy structures contributing to the first feature (peak or inflection point) above a threshold temperature in C_V exhibit a variety of structures for the amyloid-forming monomers (Figure 3). Selected atoms in the hexapeptides are used to define end-to-end distance and dihedral value parameters. These parameters are useful to classify the variety of conformations that occur for the amyloid monomers (Figure 4a). The low-temperature feature in C_V for amyloid monomers usually corresponds to a transition between structures with different backbone conformations with different main chain or side chain interactions. A similar heat capacity analysis for dimers shows another pattern in the end-to-end distance versus dihedral correlation plots. The low-energy structures that contribute to the low-temperature C_V feature exhibit larger end-to-end distances and larger positive dihedrals for the amyloid-forming sequences compared to the controls (Figure 4b).

Table 1. List of hexapeptides for which both the monomer and dimer energy landscapes have been explored [10,23,38–40]. The UniProt [41] sequence numbers for the hexapeptides are presented in Table S1 (Supplementary Materials).

Disease	Protein	Amyloid	Control
Diabetes mellitus	Amylin (hIAPP)	NFGAIL [42]	NLGPVL, LIAGFN [42], NAGAIL [43]
Alzheimer's	Tau	VQIVYK [44]	VQIVEK, NAEVYK [44]
-	De novo designed	STVIIE [45]	SPVIIE, STVIIP, STVVIE [23,45]
-	Insulin A chain	LYQLEN [10,39]	YQLENY [23,39]
Dialysis-related amyloidosis	β_2-microglobulin	LLYYTE [23,39]	YYTEFT [23]

Table 2. List of hexapeptides for which only the monomer energy landscapes have been explored [10,38]. The UniProt [41] sequence numbers for the hexapeptides are presented in Table S1 (Supplementary Materials).

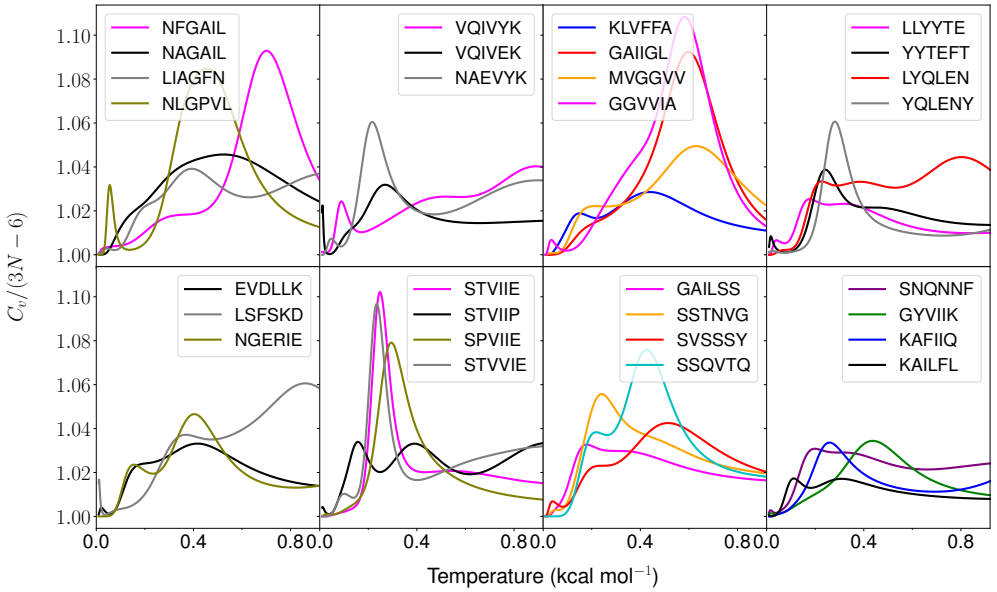
Protein	Amyloid
Aβ-A4	KLVFFA [11], GAIIGL, GGVVIA, MVGGVV [10]
Amylin	GAILSS, SSTNVG
Apolipoprotein E	SSQVTQ
Major prion protein	SNQNNF [10]
Ig κ chain	SVSSSY [38]
Serum amyloid P	GYVIIK
Protein	Control
β_2-microglobulin	EVDLLK, LSFSKD, NGERIE
Waltz-DB [46]	KAFIIQ, KAILFL

We also investigated the correlation between the aggregation propensity (calculated using Aggrescan [47] software (http://bioinf.uab.es/aggrescan/)) and the temperature at which the first feature of interest occurs in the C_V plot (Figure 5). While there is little correlation between these two quantities, further analysis suggests that a higher aggregation propensity correlates with a lower temperature for the C_V feature in sequences from the same protein. Extrinsic factors, such as pH, buffer conditions, and protein concentration [48], are known to affect amyloid formation. The present study aims to complement these results by probing the thermodynamics of sequence-specific properties for amyloid-forming peptides. Water also contributes to interactions, but solvent dynamics are difficult to visualise and quantify [49], and here, water is modelled using an implicit solvent.

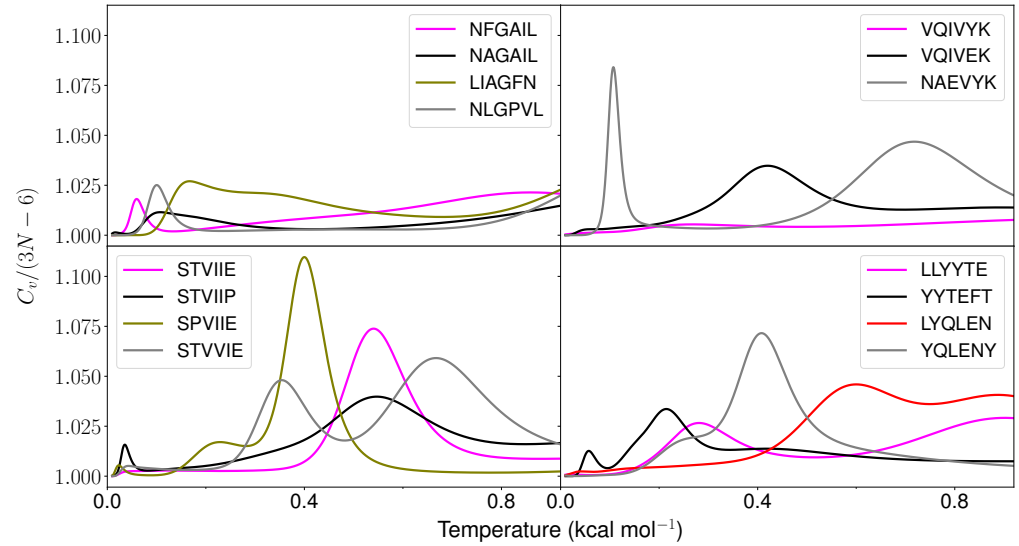
2. Results and Discussion

Various hexapeptides from naturally occurring proteins have been found to be important contributors of protein aggregation, and these segments can themselves also form amyloids. For example, NFGAIL [42] and VQIVYK [44], in human islet amyloid polypeptide (hIAPP or amylin) and tau protein, respectively, form amyloids, and the aggregates of these proteins are found in type II diabetes and Alzheimer's, respectively. Here, we explore the potential energy landscapes of amyloid-forming hexapeptides and control sequences to investigate if there is any incipient signature for this behaviour in the heat capacity, which tells us about the competition between alternative low-lying conformations that differ significantly in their enthalpy and entropy. The hexapeptides NFGAIL, VQIVYK, and STVIIE were chosen because several control studies for these peptides exist in the literature (Tables 1 and 2). The amyloid-forming peptides that contain leucine as the first residue, such as LYQLEN and LLYYTE, were chosen and compared with similar control sequences YQLENY and YYTEFT to understand the importance of leucine as the initial residue. While amyloid fibrils are hallmarks of various diseases, a previous study showed that the amyloid fibrils of hexapeptides can attenuate neuroinflammation [38]. The hexapeptides that form fibrils at a neutral pH were taken from taken from that report. These species include the cationic peptides, VQIVYK, GYVIIK, and KLVFFA; the nonionisable polar peptides SNQNNF, SSQVTQ, SSTNVG, and SVSSSY; and the nonionisable hydrophobic peptides GAIIGL, MVGGVV, GGVVIA, and GAILSS. For further validation of our results, we chose a few more control hexapeptides from β_2-microglobulin, including EVDLLK, LSFSKD, and NGERIE, while the lysine and phenylalanine containing control peptides KAFIIQ and KAILFL were chosen from Waltz-DB [46], for comparison with the amyloid-forming KLVFFA peptide.

The heat capacity analysis was performed on the converged landscapes (Figure 1). Both the control and amyloid-forming hexapeptides were found to exhibit features at low temperatures. The temperatures for all the features are listed in Tables S2–S4 in the Supplementary Materials. Interestingly, the structures with varying backbone conformations were found to contribute significantly to the low-temperature feature in amyloidogenic peptides. Note that the feature (peak/inflection point) of interest was taken as the first feature that occurred above $k_B T = 0.086\,\text{kcal mol}^{-1}$ in C_V plot. This threshold was employed because the presence of specific residues (isoleucine, valine, leucine and tyrosine) leads to low-lying structures separated by relatively small barriers. These barriers occur due to the presence of different conformations for the side chains and different rotamers for tyrosine. The N-terminal and C-terminal caps and the residues at the ends of the peptides are relatively free to orient in different directions and establish H-bonding within the peptide in different ways. Such structures are also separated by relatively small barriers. Although these structures do not differ in the main chain conformations, they may give rise to features (peaks or inflection points) below $k_B T = 0.086\,\text{kcal mol}^{-1}$ in C_V. Hence, the low-temperature peak of interest that represents a transition between structures with different main chain conformations is the first peak that occurs above the threshold temperature.



(**a**) Monomers



(**b**) Dimers

Figure 1. Heat capacity as a function of temperature ($k_B T$) for various amyloid-forming and control hexapeptide sequences.

The energy landscapes for monomers, visualised using disconnectivity graphs [50,51], can be multifunneled for both amyloid-forming and control hexapeptides (Figure 2). However, the structures lying at the bottom of funnels separated by significant barriers differ significantly for the amyloids and controls (Figures S2–S31 in Supplementary Materials). The low-energy minima separated by large barriers interconvert via the breaking of H-bonds between main chain and side chain atoms, opening of the peptide backbone, and refolding, which results in a different backbone conformation.

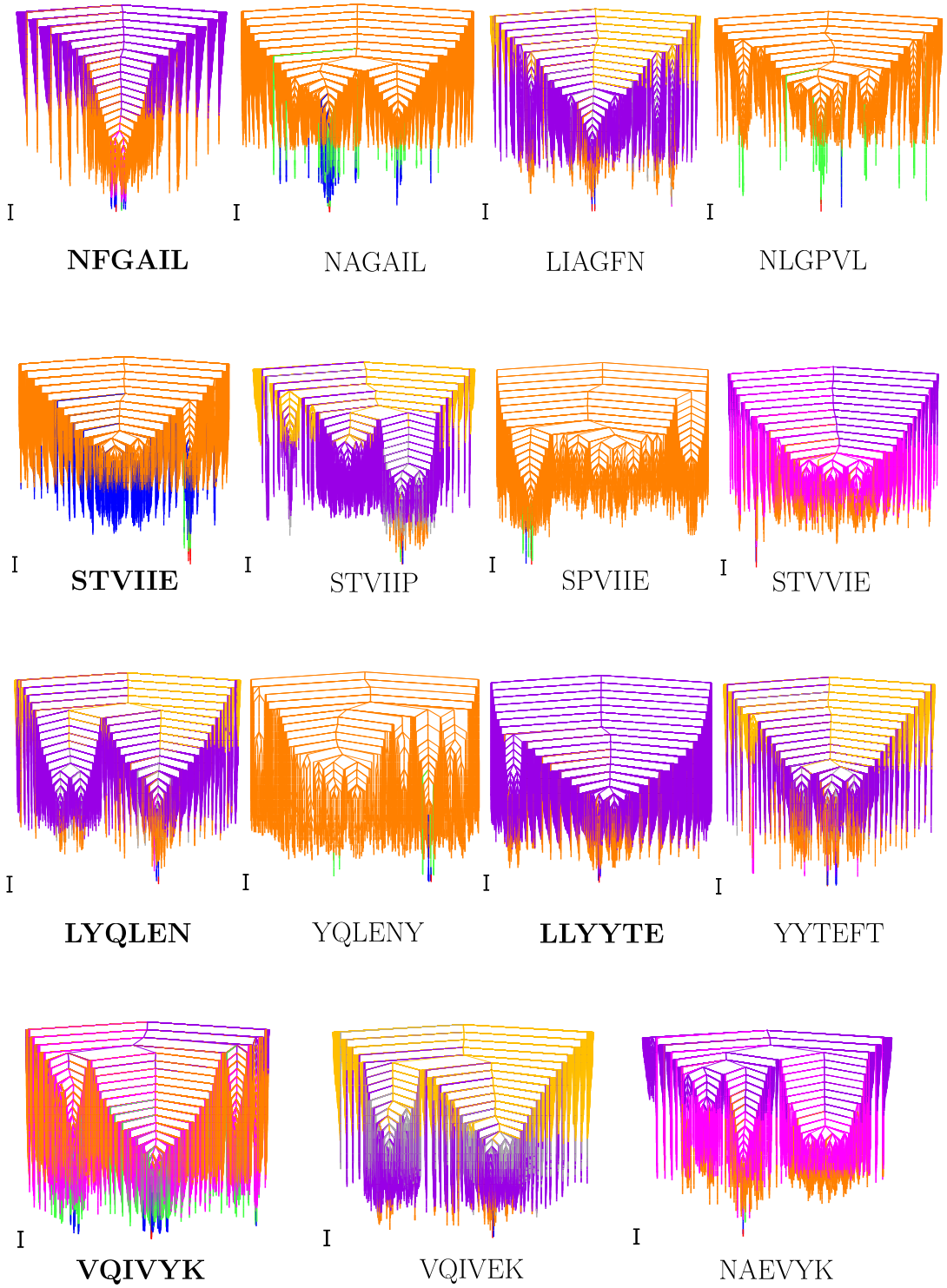
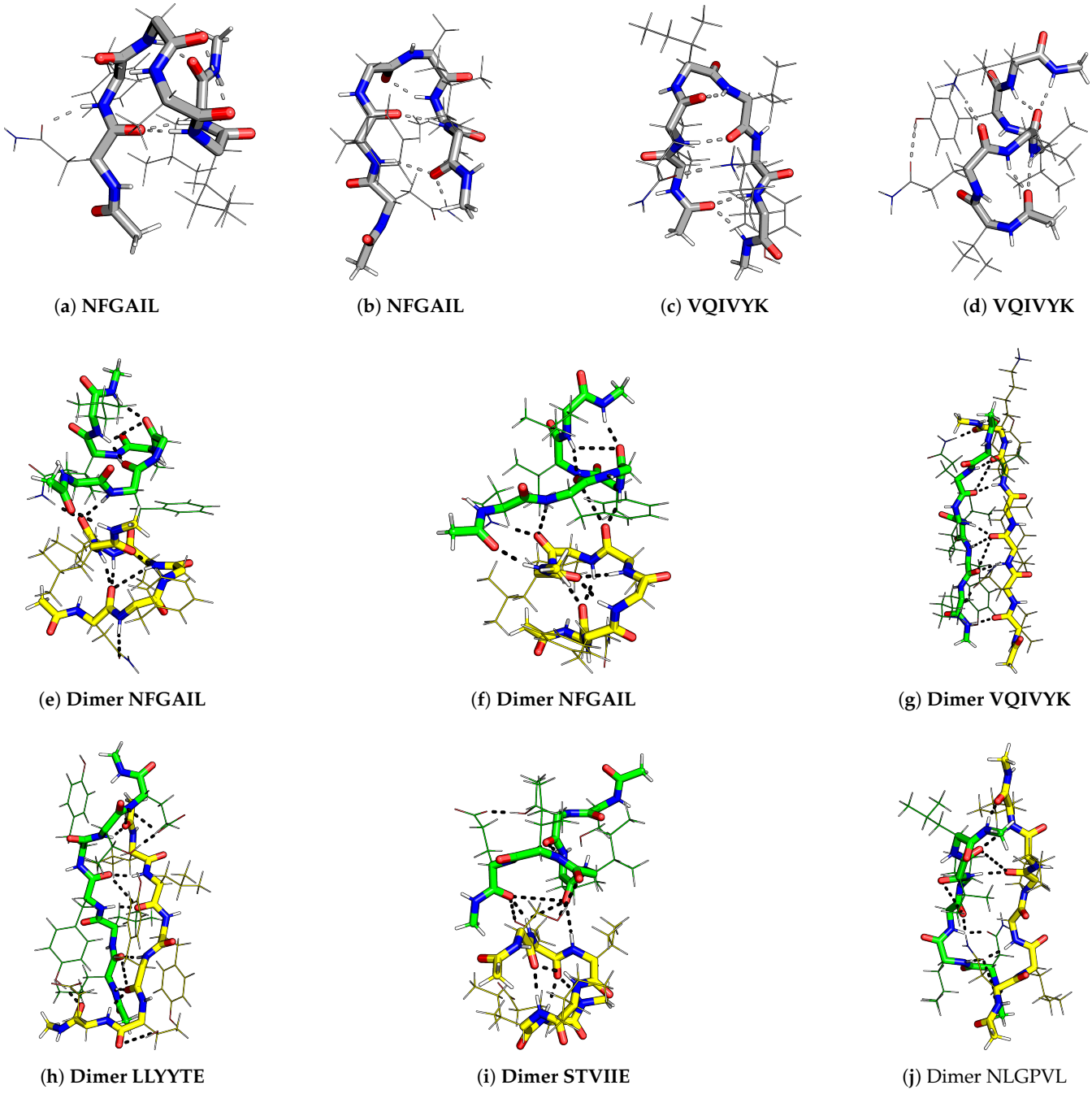


Figure 2. Disconnectivity graphs for amyloid (marked in bold) and control hexapeptide monomers. Local minima contributing to C_V features (peaks/inflection points from low to high temperature are presented in the Supplementary Materials, Figures S2–S31) are represented by red to blue (feature 1), green to orange (feature 2), pink to purple (feature 3), and grey to yellow (feature 4). The scalebar represents 1 kcal mol^{-1}.



(**a**) **NFGAIL** (**b**) **NFGAIL** (**c**) **VQIVYK** (**d**) **VQIVYK**

(**e**) **Dimer NFGAIL** (**f**) **Dimer NFGAIL** (**g**) **Dimer VQIVYK**

(**h**) **Dimer LLYYTE** (**i**) **Dimer STVIIE** (**j**) Dimer NLGPVL

Figure 3. Various conformations of hexapeptide monomers and dimers with large positive end-to-end distance and dihedral angles that contribute to low-temperature heat capacity feature. The amyloid-forming hexapeptides are marked in bold and the peptides in a dimer are shown in different colours for clarity.

Various conformations observed for hexapeptides appear similar to beta-hairpin (U-shaped), S-shaped (partially helical) [52], question-mark-shaped, W-shaped (almost helical), and extended Z-shaped structures, which are listed in the order of increasing end-to-end distance (Figure 3). The end-to-end distance is the distance between the N atom of the first residue and the C atom of the sixth residue in a hexapeptide. The dihedral angle is defined as the angle between the C_α atoms of the first, third, fourth, and sixth residues in the hexapeptide. Small distance and small dihedral angle parameters correspond to hairpin-like structures, and large distance and large dihedral angle correspond to helical structures. Even larger distances signify extended monomer structures. These distance and dihedral angle parameters were calculated for the structures contributing to the low-

temperature C_V feature, and we examined the correlations with the amyloid formation propensity. The quantitative correlation between the exact temperature at which the C_V feature occurs and the aggregation propensity of the peptide was also investigated using a different threshold temperature ($k_BT = 0.076\,\text{kcal mol}^{-1}$).

2.1. Monomer Heat Capacity

Interesting patterns were found on analysing the funnels that contained structures contributing to the first C_V feature (peak or inflection point) above $k_BT = 0.086\,\text{kcal mol}^{-1}$ (Figure 1a). In general, for amyloid-forming hexapeptide sequences, such as NFGAIL, VQIVYK, STVIIE, LYQLEN, and KLVFFA, several significantly different main chain conformations occur in different funnels (Figure 3). Both the helical and hairpin structures consistently appear together in the landscapes for amyloid hexapeptides. In contrast, for the controls (NAGAIL, VQIVEK, NAEVYK, YYTEFT, EVDLLK, LSFSKD, NGERIE, STVIIP, STVVIE, YQLENY, KAFIIQ, and KAILFL), only one type of (or a set of states with a similar secondary structure) main chain conformation occurs at the funnel bottom. This result is also evident from the greater spread of points for amyloid-forming sequences compared to the respective control sequences in the end-to-end distance versus dihedral correlation plot (Figure 4a). Different side chain interactions may be present in the helical and hairpin structures of amyloids. Helical structures generally have a H-bond between the side chains of two residues, while beta-hairpins have a H-bond between the side chains of another pair of residues, and one participating residue is common in establishing two different kinds of H-bond patterns in the two different main chain conformations. This situation occurs for VQIVYK, STVIIE, and LYQLEN amyloid hexapeptides. The control sequences lack such interactions between side chains. It is the residues containing hydroxyl or amide groups in their side chains that usually participate in these key interactions.

However, there are some amyloid-forming sequences that do not show a variety of low-energy backbone conformations contributing to the first C_V peak of interest. This situation occurs for GAILSS, SNQNNF (almost helical), MVGGVV (hairpin), SVSSSY, and GYVIIK (partial helical). GAIIGL and GGVVIA exhibit partial helical structures. LLYYTE does not form a proper hairpin that can be classified in terms of a small end-to-end distance. These observations are also evident from the narrower range of points for these peptides in the distance versus dihedral correlation plot (Figure 4a). Similarly, the control sequences NAEVYK and SPVIIE, which at first do not appear to follow the trend shown by other control sequences, with a wider spread of points in the correlation plot, are found to follow the trend when their disconnectivity graphs are analysed. Both sequences lack contributions from helical structures. However, the occurrence of extended structures leads to points with larger distance in the correlation plot. For SPVIIE, the peak of interest is the same as the melting peak, and hence extended conformations are bound to occur at this temperature. In other words, SPVIIE does not show a low-temperature peak; it shows just a melting peak.

Various intramolecular interactions occur between amino acid side chains and the backbone in the low-energy structures that contribute to features in C_V. The residues with an -OH group, such as serine, threonine, and tyrosine, can establish H-bonds and interact via H atom with the -CO group (main chain) or -COO^- group (aspartic or glutamic acid), and via the O atom with the -NH group (main chain or lysine). Similarly, the presence of an amide group in asparagine and glutamine allows the same residue to establish two different types of H-bonds, i.e., one in which the -CO group interacts and another in which the -NH group interacts. The specific interactions found between amino acid pairs in some of the low-energy structures for the hexapeptide monomers are summarised in Table 3.

Table 3. Intramolecular interactions between pairs of amino acid residues in some of the low-energy structures of hexapeptide monomers as visualised from the disconnectivity graphs (Figures S2–S31 in Supplementary Materials). The symbols X and U in parentheses represent the interactions present in helical/partial helical and hairpin structures, respectively. The aromatic rings (F/Y) may interact via a T-shaped or an offset stacked geometry. In SNQNNF, we also found three residues interacting simultaneously in the low-energy structures for the monomer.

Amyloid	Interactions	Control	Interactions
NFGAIL	N–F, F–I/L	LIAGFN	N–F
VQIVYK	I–Y, Q–K (U), Q–Y (X)	VQIVEK	K–Q/E, Q–E
STVIIE	S–E (U), T–E (X)	SPVIIE	S–E
LYQLEN	Y–Q, Q–E (U), N–E (X)	YQLENY	Y–Y, Q–E, Y–Q, Y–Q–Y
LLYYTE	Y–Y	YYTEFT	T–E, E–Y
KLVFFA	F–F	EVDLLK	K–D/E
SSQVTQ	S–T/Q, Q–Q	NAEVYK	E–Y/K
SNQNNF	S–N, N–N, Q–N, N–Q–N	STVVIE	S–E, T–E
SSTNVG	S–N/S	LSFSKD	K–D, L–F, S–D
SVSSSY	S–Y/S	NGERIE	E–R

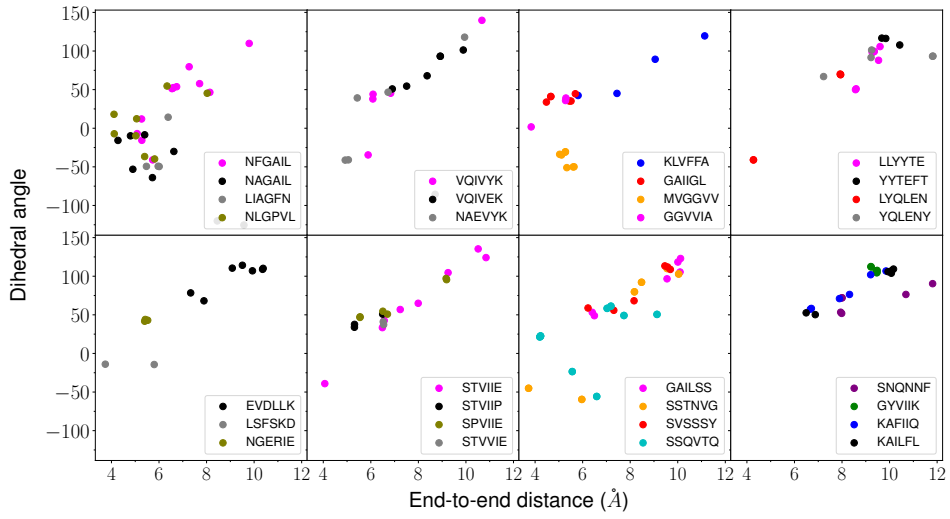
2.2. Dimer Heat Capacity

Basin-hopping [53] with rigid body moves and subsequent all-atom relaxation was used to sample the energy landscape for dimers. The low-temperature heat capacity feature was taken as the first peak or inflection point above 0.086 kcal mol^{-1}, as for the monomers (Figure 1b). The structures within 2.6 kcal mol^{-1} of the global minimum were used in the end-to-end distance versus dihedral correlation plot. Interestingly, for amyloid hexapeptides, the structures with relatively large end-to-end distances and large positive dihedrals contribute to the C_V feature of interest (Figure 4b). For VQIVYK, very large negative dihedrals are also observed along with very large positive dihedrals. When the dihedral angles are close to 180 degrees (positive or negative), the corresponding structures are similar. However, for LYQLEN and YQLENY, the points in the correlation plot are close to each other. We note that LYQLEN forms amyloid at a low pH [38], whereas the peptides were studied at a neutral pH in the current study. NLGPVL is a control peptide that exhibited extended structures contributing to the C_V peak of interest. However, further structural analysis revealed that the interstrand separation between monomers was larger compared to other amyloid-forming dimers exhibiting extended conformations (Figure 3). This result may be rationalised by taking into account the contribution from the L residue in the second position and the P residue in NLGPVL. We suggest that these residues may not allow the strands to interact very closely and hence play a role in preventing the aggregation of this peptide.

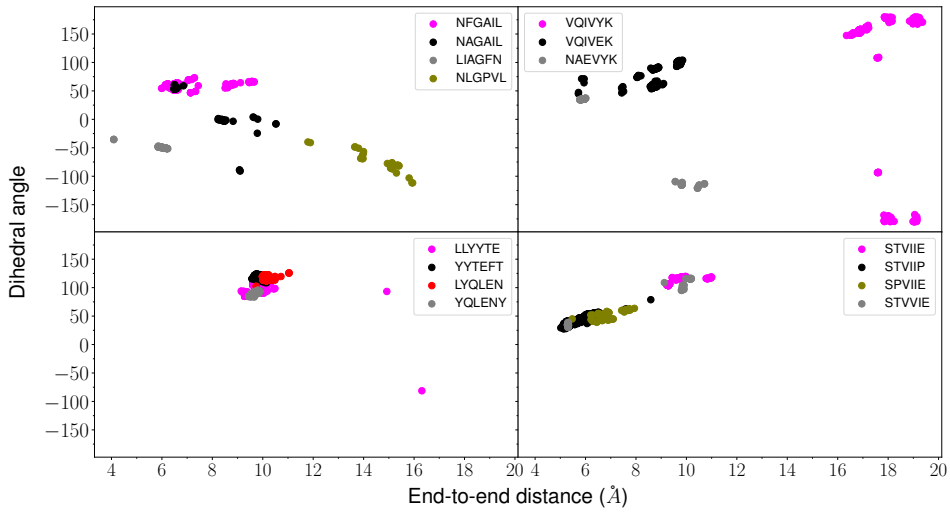
2.3. Correlation between Heat Capacity and Amyloid Formation Predictors

The correlation plot (Figure 5) between the temperature at which the low-temperature feature occurs in the C_V plot and the propensity for amyloid formation, as predicted using Aggrescan [47], shows that there is little overall correlation between the two quantities. However, further analysis reveals that the amyloid and control sequences derived from the same protein can probably be distinguished by comparing the temperature at which the first feature (excluding melting) occurs in the heat capacity plot between 0.076 and 0.300 kcal mol^{-1}. The amyloid-forming sequences exhibit features at a lower temperature compared to sequences with a lower propensity for amyloid formation and the respective control sequences. This trend is evident when we compare the sequences within the sets derived from the same protein, such as tau (VQIVYK, VQIVEK, and NAEVYK), amylin (NFGAIL and NAGAIL), β_2-microglobulin (LLYYTE, YYTEFT and EVDLLK), and A-β (KLVFFA and MVGGVV). The control sequences KAFIIQ, NLGPVL, and YQLENY lacked

such low-temperature features. Another predictor that can be useful to estimate amyloid formation propensity is CamSol [54]. CamSol is designed to predict the intrinsic solubility of proteins. Aggregation-prone sequences exhibit lower solubility. Heptapeptides are the minimal sequences for which solubility can be obtained using CamSol. To obtain an approximate idea of solubility for the capped hexapeptides, we added alanine residues at the termini of the peptide sequences. As expected, the solubility exhibited a strong negative correlation with the aggregation propensity predicted using Aggrescan (Figure S1 in Supplementary Materials). Hence, the low-temperature heat capacity feature is positively correlated with intrinsic solubility for peptide sequences occurring in the same context, i.e., within the same protein (Figure S1 in Supplementary Materials). As for Aggrescan, there is little overall correlation between the two quantities, intrinsic solubility and low-temperature C_V feature. We note that the CamSol and Aggrescan predictors may have some intrinsic limitations for such small sequences. Both of them associate the control sequence KAILFL with high aggregation propensity, and the amyloid sequences with two or more serine residues are predicted to have high solubility and lower aggregation propensity. Overall, low-temperature heat capacity features for proteins do not exhibit a strong correlation with aggregation propensity. However, they may be useful to compare the properties of short sequences occurring within the same protein, i.e., within the same context.



(**a**) Monomers



(**b**) Dimers

Figure 4. Correlation plots between end-to-end distance and dihedral angles for low-energy structures contributing to first heat capacity peak of interest.

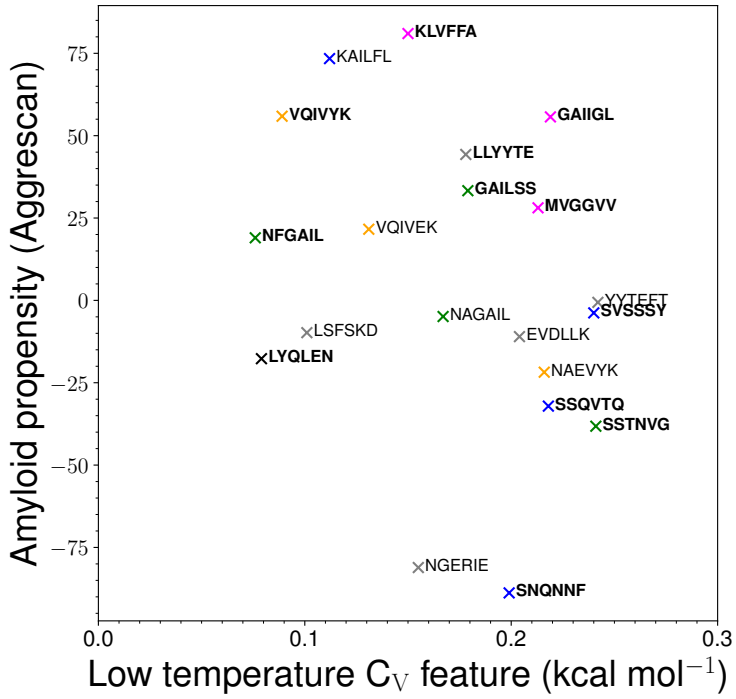


Figure 5. Correlation plot between the temperature ($k_B T$) at which the first low-temperature feature is observed (between 0.076 and 0.300 kcal mol^{-1}) in the C_V plot for the monomer and the propensity for amyloid formation predicted using the Aggrescan software [47]. The colours correspond to the proteins from which the sequences are taken. Orange, green, magenta, and grey correspond to tau, amylin, A-β, and β_2-microglobulin, respectively. Blue represents miscellaneous peptides taken from various proteins. Black represents a peptide which is amyloidogenic at a different pH. Several peptides are excluded from the above plot. The excluded peptides include GYVIIK, KAFIIQ, NLGPVL, and YQLENY, where the first C_V feature is simply the melting peak; GGVVIA, which has such a feature occurring above 0.3 kcal mol^{-1}; the LIAGFN control sequence, for which the existing predictor fail to classify it differently from its reverse (NFGAIL) amyloid-forming sequence; and the de novo designed peptides STVIIE, STVIIP, SPVIIE, and STVVIE, which do not derive from naturally occurring amyloid-forming proteins.

3. Materials and Methods

The hexapeptides were modelled using the FF99IDPs [55] force field, which is a modified version of FF99SBILDN [56] within AMBER [57,58]. This choice was made to gain better insight into the energy landscapes of various sequences using the same potential and compare them with our earlier study [37] using the same force field. Our previous research [37] showed that the structures that contribute to low-temperature C_V peaks for different AMBER force fields are similar. The N- and C-terminals of the peptide were methylated and methylamidated, respectively. The topology file was symmetrised [59] to account correctly for permutational isomers. Water was modelled using implicit solvation (igb = 8) along with a 0.1 M monovalent ion concentration [57,58].

The initial landscape exploration was performed using basin-hopping parallel tempering (BHPT) [53,60,61], implemented within the global optimisation program GMIN [62] interfaced with AMBER. Similar to our previous study [37] on monomers, 16 replicas were used for BHPT, with temperatures exponentially distributed between 300 and 575 K. For dimers, the potential energy landscapes were explored using a combination of rigid body, Cartesian, and group rotation moves [63–65]. Each monomer was used to define a rigid body. The two monomers were expanded radially from the centre of the coordinates, rotated within the angle–axis framework [66], and translated, with rigid body moves performed after every 111 Cartesian moves. The thousand lowest energy structures were saved and converged tighter to a root-mean-square convergence criterion for the gradient

of 10^{-7} kcal mol^{-1} A^{-1}. A total of 600,000–800,000 basin-hopping steps were performed for each peptide dimer. This total involves different runs with different starting structures, step size, and frequency of rigid body moves. For dimers, the sampling was monitored using the convergence of low-temperature heat capacity peaks.

For monomers, discrete path sampling [67] was employed to obtain pathways for interconversion between local minima and the global minimum. Each multistep pathway was composed of minimum-transition state-minimum triples [68,69]. Several geometry optimisation algorithms were employed to obtain these pathways [70], including the doubly nudged [71,72] elastic band [73–76] algorithm, hybrid eigenvector-following [77–80], minimisation with a limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm [81,82], and Dijkstra's shortest-path algorithm [83]. These tools were implemented within the OPTIM [84] program. The PATHSAMPLE [85] program was used to expand the stationary point database, including a strategy to remove unphysical barriers connecting various local minima [86]. The convergence of the database was monitored using disconnectivity graphs and the convergence of low-temperature heat capacity peaks.

Disconnectivity graphs [50,51] provide an overview of the landscape organisation. These graphs preserve the information about the highest-energy barrier that needs to be overcome to interconvert a pair of connected local minima. The vertical axis represents the potential or free energy, and the nodes along the vertical axis define superbasins. The minima within a superbasin can interconvert by overcoming a barrier that is less than or equal to the threshold energy associated with the corresponding node. The branches terminate at the potential or free energy of the individual local minimum.

The heat capacity of the peptides was estimated using the harmonic superposition approximation [87–92]. The partition function of each local minima was obtained using normal mode analysis and the total partition function was the sum of the partition functions of all the local minima. Each peak in the heat capacity involves contributions from minima with negative and positive occupation probability derivatives with respect to temperature [93]. The principal contributions (98%) for each heat capacity peak were visualised in disconnectivity graphs using different-coloured branches for each set of minima (Figures S2–S31 in Supplementary Materials). The geometric parameters (end-to-end distance and dihedral angle) used to distinguish different structures were calculated using the CPPTRAJ program within AMBER [94].

4. Conclusions

The heat capacity analysis of peptide monomers and dimers may provide insight into collective phenomena, such as the aggregation and phase separation of proteins. We found that a variety of low-energy structures with different backbone conformations contributed to the low-temperature heat capacity feature for amyloid-forming hexapeptide monomers. For control sequences, the C_V peak of interest did not correspond to such a diverse set of structures. The structural competition can be diagnosed via a combination of geometric parameters, such as end-to-end distance and dihedral angle. The heat capacity analysis for dimer conformations revealed that the amyloid-forming hexapeptides preferentially contributed to the peak of interest via low-energy conformations with large end-to-end distance and large positive dihedrals. The exact temperature at which the C_V feature occurred was lower for the peptides with higher propensity for amyloid formation compared to control sequences derived from the same protein. However, this trend does not hold if we compare amyloidogenic hexapeptides from one protein and control sequences from a different protein.

We do not expect to extract a universal aggregating propensity from monomer and dimer properties. However, the analysis of low-temperature heat capacity features revealed new opportunities for investigating and predicting collective behaviour. The distinctions we have identified for amyloid-forming hexapeptides, compared to controls with very similar sequences, suggest that further investigation could be fruitful. For example, the species responsible for aggregation could involve higher free energy minima (excitations [52]) in

the monomer landscape, which may contribute to heat capacity features. The fact that the current analysis could distinguish amyloid and control sequences with the same amino acid residues in reverse order, as for NFGAIL and LIAGFN, further illustrates how the energy landscape framework can be used to understand the context-dependent properties of amino acid residues in proteins. In the future, we will investigate the heat capacities of oligomers using coarse-grained potentials.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms241310613/s1.

Author Contributions: N. conceived the idea and designed the study with the help of D.J.W. N. performed the simulations and wrote the first draft. Both the authors analysed and interpreted the data, and edited the final version of the draft. D.J.W. supervised the project and developed all the original energy landscape software. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) (D.J.W., grant number EP/N035003/1); the Cambridge Commonwealth, European and International Trust; the Allen, Meek and Read Fund; the Santander fund, St Edmund's College, University of Cambridge; and the Trinity-Henry Barlow Honorary Award (Nicy).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The discrete path sampling databases are available at https://doi.org/10.17863/CAM.96759 (accessed on 25 May 2023) [95]. The computational protocol for creating the database is available as a tutorial on https://github.com/nicy-nicy/peptide-energy-landscape-exploration (accessed on 18 May 2023) and the scripts to analyse the database are available on https://github.com/nicy-nicy/energy-landscape-cv-analysis (accessed on 18 May 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C.M.; Stefani, M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **2002**, *416*, 507–511. [CrossRef] [PubMed]
2. Sunde, M.; Blake, C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Protein Chem.* **1997**, *50*, 123–159. [PubMed]
3. Booth, D.R.; Sunde, M.; Bellotti, V.; Robinson, C.V.; Hutchinson, W.L.; Fraser, P.E.; Hawkins, P.N.; Dobson, C.M.; Radford, S.E.; Blake, C.C.; et al. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **1997**, *385*, 787–793. [CrossRef] [PubMed]
4. Fändrich, M.; Fletcher, M.A.; Dobson, C.M. Amyloid fibrils from muscle myoglobin. *Nature* **2001**, *410*, 165–166. [CrossRef] [PubMed]
5. Krebs, M.R.; Morozova-Roche, L.A.; Daniel, K.; Robinson, C.V.; Dobson, C.M. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci.* **2004**, *13*, 1933–1938. [CrossRef]
6. Fernandez-Escamilla, A.M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **2004**, *22*, 1302–1306. [CrossRef]
7. Balbirnie, M.; Grothe, R.; Eisenberg, D.S. An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β-sheet structure for amyloid. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2375–2380. [CrossRef]
8. Ventura, S.; Zurdo, J.; Narayanan, S.; Parreño, M.; Mangues, R.; Reif, B.; Chiti, F.; Giannoni, E.; Dobson, C.M.; Aviles, F.X.; et al. Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7258–7263. [CrossRef]
9. Nelson, R.; Sawaya, M.R.; Balbirnie, M.; Madsen, A.Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-β spine of amyloid-like fibrils. *Nature* **2005**, *435*, 773–778. [CrossRef]
10. Sawaya, M.R.; Sambashivan, S.; Nelson, R.; Ivanova, M.I.; Sievers, S.A.; Apostol, M.I.; Thompson, M.J.; Balbirnie, M.; Wiltzius, J.J.; McFarlane, H.T.; et al. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **2007**, *447*, 453–457. [CrossRef]
11. Tjernberg, L.O.; Callaway, D.J.; Tjernberg, A.; Hahne, S.; Lilliehöök, C.; Terenius, L.; Thyberg, J.; Nordstedt, C. A molecular model of Alzheimer amyloid β-peptide fibril formation. *J. Biol. Chem.* **1999**, *274*, 12619–12625. [CrossRef] [PubMed]

12. Zanuy, D.; Haspel, N.; Tsai, H.H.G.; Ma, B.; Gunasekaran, K.; Wolfson, H.J.; Nussinov, R. Side chain interactions determine the amyloid organization: A single layer β-sheet molecular structure of the calcitonin peptide segment 15–19. *Phys. Biol.* **2004**, *1*, 89. [CrossRef] [PubMed]
13. Knowles, T.P.; Fitzpatrick, A.W.; Meehan, S.; Mott, H.R.; Vendruscolo, M.; Dobson, C.M.; Welland, M.E. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **2007**, *318*, 1900–1903. [CrossRef]
14. Wang, X.; Chapman, M.R. Sequence determinants of bacterial amyloid formation. *J. Mol. Biol.* **2008**, *380*, 570–580. [CrossRef]
15. Gazit, E. A possible role for π-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **2002**, *16*, 77–83. [CrossRef] [PubMed]
16. Tjernberg, L.; Hosia, W.; Bark, N.; Thyberg, J.; Johansson, J. Charge attraction and β propensity are necessary for amyloid fibril formation from tetrapeptides. *J. Biol. Chem.* **2002**, *277*, 43243–43246. [CrossRef]
17. Marshall, K.E.; Morris, K.L.; Charlton, D.; O'Reilly, N.; Lewis, L.; Walden, H.; Serpell, L.C. Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability. *Biochemistry* **2011**, *50*, 2061–2071. [CrossRef]
18. Berhanu, W.M.; Masunov, A.E. Can molecular dynamics simulations assist in design of specific inhibitors and imaging agents of amyloid aggregation? Structure, stability and free energy predictions for amyloid oligomers of VQIVYK, MVGGVV and LYQLEN. *J. Mol. Model.* **2011**, *17*, 2423–2442. [CrossRef]
19. Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C.M. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **2003**, *424*, 805–808. [CrossRef]
20. Tartaglia, G.G.; Cavalli, A.; Pellarin, R.; Caflisch, A. The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates. *Protein Sci.* **2004**, *13*, 1939–1941. [CrossRef]
21. Esler, W.P.; Stimson, E.R.; Fishman, J.B.; Ghilardi, J.R.; Vinters, H.V.; Mantyh, P.W.; Maggio, J.E. Stereochemical specificity of Alzheimer's disease β-peptide assembly. *Biopolymers* **1999**, *49*, 505–514. [CrossRef]
22. Pawar, A.P.; DuBay, K.F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C.M. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **2005**, *350*, 379–392.
23. Thompson, M.J.; Sievers, S.A.; Karanicolas, J.; Ivanova, M.I.; Baker, D.; Eisenberg, D. The 3D profile method for identifying fibril-forming segments of proteins. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4074–4078. [CrossRef] [PubMed]
24. Tartaglia, G.G.; Vendruscolo, M. The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev.* **2008**, *37*, 1395–1401.
25. Goldschmidt, L.; Teng, P.K.; Riek, R.; Eisenberg, D. Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 3487–3492. [CrossRef]
26. Nekooki-Machida, Y.; Kurosawa, M.; Nukina, N.; Ito, K.; Oda, T.; Tanaka, M. Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9679–9684. [CrossRef] [PubMed]
27. Eichner, T.; Radford, S.E. A diversity of assembly mechanisms of a generic amyloid fold. *Mol. Cell* **2011**, *43*, 8–18. [CrossRef]
28. Marshall, K.E.; Hicks, M.R.; Williams, T.L.; Hoffmann, S.V.; Rodger, A.; Dafforn, T.R.; Serpell, L.C. Characterizing the assembly of the Sup35 yeast prion fragment, GNNQQNY: Structural changes accompany a fiber-to-crystal switch. *Biophys. J.* **2010**, *98*, 330–338. [CrossRef] [PubMed]
29. DuBay, K.F.; Pawar, A.P.; Chiti, F.; Zurdo, J.; Dobson, C.M.; Vendruscolo, M. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.* **2004**, *341*, 1317–1326. [CrossRef]
30. Buell, A.K. Stability matters, too–the thermodynamics of amyloid fibril formation. *Chem. Sci.* **2022**, *13*, 10177–10192. [CrossRef]
31. Gómez, J.; Hilser, V.J.; Xie, D.; Freire, E. The heat capacity of proteins. *Proteins Struct. Funct. Bioinform.* **1995**, *22*, 404–412. [CrossRef] [PubMed]
32. Cooper, A.; Johnson, C.M.; Lakey, J.H.; Nöllmann, M. Heat does not come in different colours: Entropy–enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophys. Chem.* **2001**, *93*, 215–230. [CrossRef] [PubMed]
33. Loladze, V.V.; Ermolenko, D.N.; Makhatadze, G.I. Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* **2001**, *10*, 1343–1352. [CrossRef] [PubMed]
34. Kardos, J.; Yamamoto, K.; Hasegawa, K.; Naiki, H.; Goto, Y. Direct measurement of the thermodynamic parameters of amyloid formation by isothermal titration calorimetry. *J. Biol. Chem.* **2004**, *279*, 55308–55314. [CrossRef]
35. Jeppesen, M.D.; Hein, K.; Nissen, P.; Westh, P.; Otzen, D.E. A thermodynamic analysis of fibrillar polymorphism. *Biophys. Chem.* **2010**, *149*, 40–46. [CrossRef]
36. Rana, N.; Kodirov, R.; Shakya, A.; King, J.T. Protein unfolding thermodynamics predict multicomponent phase behavior. *bioRxiv* **2023**. [CrossRef]
37. Nicy.; Joseph, J.A.; Collepardo-Guevara, R.; Wales, D.J. Energy landscapes and heat capacity signatures for peptides correlate with phase separation propensity. *bioRxiv* **2023**. [CrossRef]
38. Kurnellas, M.P.; Adams, C.M.; Sobel, R.A.; Steinman, L.; Rothbard, J.B. Amyloid fibrils composed of hexameric peptides attenuate neuroinflammation. *Sci. Transl. Med.* **2013**, *5*, 179ra42. [CrossRef]
39. Ivanova, M.I.; Thompson, M.J.; Eisenberg, D. A systematic screen of β2-microglobulin and insulin for amyloid-like segments. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4079–4082. [CrossRef]
40. Makin, O.S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L.C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 315–320. [CrossRef]

41. Consortium, T.U. UniProt: The Universal Protein knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523–D531. [CrossRef] [PubMed]
42. Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M.L.; Voelter, W.; Brunner, H.; Kapurniotu, A. Identification of a penta-and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. *J. Mol. Biol.* **2000**, *295*, 1055–1071. [CrossRef]
43. Azriel, R.; Gazit, E. Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. *J. Biol. Chem.* **2001**, *276*, 34156–34161. [CrossRef] [PubMed]
44. Von Bergen, M.; Friedhoff, P.; Biernat, J.; Heberle, J.; Mandelkow, E.M.; Mandelkow, E. Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (^{306}VQIVYK^{311}) forming β structure. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5129–5134. [CrossRef]
45. López de la Paz, M.; Serrano, L. Sequence determinants of amyloid fibril formation. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 87–92. [CrossRef] [PubMed]
46. Louros, N.; Konstantoulea, K.; De Vleeschouwer, M.; Ramakers, M.; Schymkowitz, J.; Rousseau, F. WALTZ-DB 2.0: An updated database containing structural information of experimentally determined amyloid-forming peptides. *Nucleic Acids Res.* **2020**, *48*, D389–D393. [CrossRef]
47. Conchillo-Solé, O.; de Groot, N.S.; Avilés, F.X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: A server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinform.* **2007**, *8*, 65. [CrossRef] [PubMed]
48. Ciryam, P.; Tartaglia, G.G.; Morimoto, R.I.; Dobson, C.M.; Vendruscolo, M. Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Rep.* **2013**, *5*, 781–790. [CrossRef]
49. Johnson, C.M. Differential scanning calorimetry as a tool for protein folding and stability. *Arch. Biochem. Biophys.* **2013**, *531*, 100–109. [CrossRef]
50. Becker, O.M.; Karplus, M. The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *J. Chem. Phys.* **1997**, *106*, 1495–1517. [CrossRef]
51. Wales, D.J.; Miller, M.A.; Walsh, T.R. Archetypal energy landscapes. *Nature* **1998**, *394*, 758–760. [CrossRef]
52. Chakraborty, D.; Straub, J.E.; Thirumalai, D. Energy landscapes of Aβ monomers are sculpted in accordance with Ostwald's rule of stages. *Sci. Adv.* **2023**, *9*, eadd6921. [CrossRef] [PubMed]
53. Wales, D.J.; Doye, J.P.K. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **1997**, *101*, 5111–5116. [CrossRef]
54. Sormanni, P.; Aprile, F.A.; Vendruscolo, M. The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **2015**, *427*, 478–490. [CrossRef] [PubMed]
55. Wang, W.; Ye, W.; Jiang, C.; Luo, R.; Chen, H.F. New force field on modeling intrinsically disordered proteins. *Chem. Biol. Drug Des.* **2014**, *84*, 253–269. [CrossRef]
56. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinform.* **2010**, *78*, 1950–1958. [CrossRef]
57. Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [CrossRef]
58. Case, D.A.; Duke, R.E.; Walker, R.C.; Skrynnikov, N.R.; Cheatham T.E., III; Mikhailovskii, O.; Simmerling, C.; Xue, Y.; Roitberg, A.; Izmailov, S.A.; et al. *AMBER 22 Reference Manual*; University of California: Los Angeles, CA, USA, 2022.
59. Malolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S.N.; Wales, D.J. Symmetrization of the AMBER and CHARMM Force Fields. *J. Comput. Chem.* **2010**, *31*, 1402–1409. [CrossRef]
60. Li, Z.; Scheraga, H.A. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6611–6615. [CrossRef] [PubMed]
61. Li, Z.; Scheraga, H.A. Structure and free energy of complex thermodynamic systems. *J. Mol. Struct. THEOCHEM* **1988**, *179*, 333–352. [CrossRef]
62. Wales, D.J. GMIN: A Program for Finding Global Minima and Calculating Thermodynamic Properties from Basin-Sampling. 2023. Available online: http://www-wales.ch.cam.ac.uk/GMIN/ (accessed on 26 January 2023)
63. Kusumaatmaja, H.; Whittleston, C.S.; Wales, D.J. A local rigid body framework for global optimization of biomolecules. *J. Chem. Theory Comput.* **2012**, *8*, 5159–5165. [CrossRef]
64. Rühle, V.; Kusumaatmaja, H.; Chakrabarti, D.; Wales, D.J. Exploring energy landscapes: Metrics, pathways, and normal-mode analysis for rigid-body molecules. *J. Chem. Theory Comput.* **2013**, *9*, 4026–4034. [CrossRef]
65. Mochizuki, K.; Whittleston, C.S.; Somani, S.; Kusumaatmaja, H.; Wales, D.J. A conformational factorisation approach for estimating the binding free energies of macromolecules. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2842–2853. [CrossRef] [PubMed]
66. Chakrabarti, D.; Wales, D.J. Simulations of rigid bodies in an angle-axis framework. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1970–1976. [CrossRef] [PubMed]
67. Wales, D.J. Discrete path sampling. *Mol. Phys.* **2002**, *100*, 3285–3305. [CrossRef]
68. Murrell, J.N.; Laidler, K.J. Symmetries of activated complexes. *Trans. Faraday Soc.* **1968**, *64*, 371–377. [CrossRef]
69. Wales, D.J. *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press: Cambridge, UK, 2003.
70. Carr, J.M.; Trygubenko, S.A.; Wales, D.J. Finding pathways between distant local minima. *J. Chem. Phys.* **2005**, *122*, 234903. [CrossRef]

71. Trygubenko, S.A.; Wales, D.J. A doubly nudged elastic band method for finding transition states. *J. Chem. Phys.* **2004**, *120*, 2082–2094. [CrossRef]
72. Sheppard, D.; Terrell, R.; Henkelman, G. Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **2008**, *128*, 134106. [CrossRef]
73. Mills, G.; Jónsson, H.; Schenter, G.K. Reversible work transition state theory: Application to dissociative adsorption of hydrogen. *Surf. Sci.* **1995**, *324*, 305–337. [CrossRef]
74. Jónsson, H.; Mills, G.; Jacobsen, K.W. Nudged elastic band method for fnding minimum energy paths of transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific: Singapore, 1998; Chapter 16, pp. 385–404.
75. Henkelman, G.; Uberuaga, B.P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904. [CrossRef]
76. Henkelman, G.; Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **2000**, *113*, 9978–9985. [CrossRef]
77. Henkelman, G.; Jónsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **1999**, *111*, 7010–7022. [CrossRef]
78. Munro, L.J.; Wales, D.J. Defect migration in crystalline silicon. *Phys. Rev. B* **1999**, *59*, 3969–3980. [CrossRef]
79. Kumeda, Y.; Wales, D.J.; Munro, L.J. Transition states and rearrangement mechanisms from hybrid eigenvector-following and density functional theory. Application to $C_{10}H_{10}$ and defect migration in crystalline silicon. *Chem. Phys. Lett.* **2001**, *341*, 185–194. [CrossRef]
80. Zeng, Y.; Xiao, P.; Henkelman, G. Unification of algorithms for minimum mode optimization. *J. Chem. Phys.* **2014**, *140*, 044115. [CrossRef] [PubMed]
81. Nocedal, J. Updating quasi-Newton matrices with limited storage. *Math. Comput.* **1980**, *35*, 773–782. [CrossRef]
82. Liu, D.C.; Nocedal, J. On the limited memory BFGS method for large scale optimization. *Math. Program.* **1989**, *45*, 503–528. [CrossRef]
83. Dijkstra, E.W. A note on two problems in connexion with graphs. *J. Numer. Math.* **1959**, *1*, 269–271. [CrossRef]
84. Wales, D.J. OPTIM: A Program for Optimising Geometries and Calculating Pathways. 2023. Available online: https://www-wales.ch.cam.ac.uk/OPTIM/ (accessed on 26 January 2023)
85. Wales, D.J. PATHSAMPLE: A Program for Generating Connected Stationary Point Databases and Extracting Global Kinetics. 2023. Available online: https://www-wales.ch.cam.ac.uk/PATHSAMPLE/ (accessed on 26 January 2023)
86. Strodel, B.; Whittleston, C.S.; Wales, D.J. Thermodynamics and kinetics of aggregation for the GNNQQNY peptide. *J. Am. Chem. Soc.* **2007**, *129*, 16005–16014. [CrossRef]
87. McGinty, D.J. Vapor phase homogeneous nucleation and the thermodynamic properties of small clusters of argon atoms. *J. Chem. Phys.* **1971**, *55*, 580–588. [CrossRef]
88. Burton, J. Vibrational frequencies and entropies of small clusters of atoms. *J. Chem. Phys.* **1972**, *56*, 3133–3138. [CrossRef]
89. Hoare, M. Structure and dynamics of simple microclusters. *Adv. Chem. Phys.* **1979**, *40*, 49–135.
90. Franke, G.; Hilf, E.; Borrmann, P. The structure of small clusters: Multiple normal-modes model. *J. Chem. Phys.* **1993**, *98*, 3496–3502. [CrossRef]
91. Wales, D.J. Coexistence in small inert gas clusters. *Mol. Phys.* **1993**, *78*, 151–171. [CrossRef]
92. Strodel, B.; Wales, D.J. Free energy surfaces from an extended harmonic superposition approach and kinetics for alanine dipeptide. *Chem. Phys. Lett.* **2008**, *466*, 105–115. [CrossRef]
93. Wales, D.J. Decoding heat capacity features from the energy landscape. *Phys. Rev. E* **2017**, *95*, 030105. [CrossRef]
94. Roe, D.R.; Cheatham III, T.E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [CrossRef]
95. Nicy; Wales, D.J. *Research Data Supporting—Energy Landscapes and Heat Capacity Signatures for Monomers and Dimers of Amyloid Forming Hexapeptides*; University of Cambridge: Cambridge, UK, 2023. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.