

Quantum chemical modelling of sequence encoded energetics and electrostatic landscapes governing early peptide association in α - synuclein and amyloid β

Oluwasesan Adegoke

`o.adeoke@dundee.ac.uk`

University of Dundee

Ojodomo J. Achadu

Teesside University

Research Article

Keywords:

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11008420/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Context Early molecular association of α -synuclein and amyloid β (A β) underlies the formation of higher order assemblies implicated in Parkinson's and Alzheimer's diseases, yet how residue order shapes the molecular properties governing these initial association events remains poorly resolved. Here, aggregation prone peptide segments were compared with composition matched scrambled controls to isolate the contribution of sequence arrangement while conserving amino acid composition. Residue rearrangement produced distinct energetic responses in the two peptide systems. Aggregation prone A β exhibited an association energy of 115.617 kJ/mol compared with 232.552 kJ/mol for its control, whereas aggregation prone α -synuclein exhibited – 172.819 kJ/mol compared with – 494.371 kJ/mol for its control. These energetic differences were accompanied by contrasting intermolecular organization: aggregation prone α -synuclein formed 166 heavy atom contacts compared with 105 for its control, whereas aggregation prone A β formed 40 compared with 81. Residue rearrangement further reorganized molecular electrostatic potential (ESP) landscapes and Hirshfeld atomic charge distributions, while sequence descriptors revealed peptide specific changes in charge patterning, hydrophobic organization and molecular compactness. Together, these results established residue order as a determinant of early peptide association and revealed that identical amino acid compositions can generate distinct energetic, structural and electrostatic characteristics through sequence dependent molecular organization.

Method Monomer and homodimer structures were optimized using the GFN2-xTB method implemented in xTB, with initial dimer relaxation using the GFN force field. Association energies were calculated from the converged GFN2-xTB electronic energies of the independently optimized monomers and corresponding dimers. DFT calculations were performed using GPAW with the PBE exchange correlation functional and PAW formalism on a real space grid with 0.20 Å spacing. ESP and Hirshfeld atomic charge analyses were obtained from the converged electronic densities. Sequence, structural and intermolecular interface descriptors were calculated using Python with the ASE, NumPy, SciPy, pandas, Matplotlib and scikit image. Open Babel was used for molecular structure conversion.

Introduction

Protein misfolding and aggregation are central molecular processes in several neurodegenerative disorders, including Parkinson's disease (PD) and Alzheimer's disease (AD)[1–4]. Protein aggregation proceeds through molecular association events that can generate soluble oligomeric and prefibrillar assemblies before the formation of mature fibrillar structures. These intermediate assemblies have been associated with cellular dysfunction, although their structures, populations and biological effects vary among proteins and experimental conditions[1, 5]. α -Synuclein and amyloid β (A β), which are closely associated with PD and AD, respectively, contain sequence regions that contribute strongly to molecular recognition and self-association [6–8]. Structural studies have shown that particular sequence segments participate in ordered fibrillar cores and that changes in residue identity or residue order can alter aggregation behaviour and the structures of the resulting assemblies [7–9]. Amino acid sequence

therefore provides a molecular basis through which the balance of intermolecular interactions governing peptide association can be modified.

Characterisation of the earliest stages of peptide association remains challenging because the participating molecular ensembles can be transient, structurally heterogeneous and continuously redistributed during aggregation. Experimental studies have provided substantial structural and kinetic information on amyloid formation but resolving the energetic contributions of individual molecular configurations within these heterogeneous ensembles remains difficult [10, 11]. Computational modelling complements these approaches by enabling chemically defined molecular structures to be examined under consistent theoretical conditions. This provides a basis for relating sequence dependent structural organisation to the energetics and electronic properties of molecular association.

Classical molecular dynamics simulations have been widely applied to protein conformational behaviour and aggregation because they provide access to molecular motions, structural rearrangements and conformational ensembles over accessible simulation timescales [12, 13]. Quantum chemical methods provide a complementary description in which molecular energies and electronic properties are derived from an explicit treatment of electronic structure. Their application to peptide systems requires a balance between molecular size, computational cost and the level of electronic structure description. Semiempirical quantum chemical methods provide this balance by enabling comparatively large molecular systems to be examined while retaining an electronic representation of bonding and noncovalent interactions [14, 15]. The Geometry, Frequency, Noncovalent extended Tight Binding, second generation (GFN2-xTB) method provides a broadly parametrised semiempirical quantum chemical framework incorporating multipole electrostatics and density dependent dispersion contributions [16]. Its formulation for molecular geometry, energetics and noncovalent interactions, together with its computational efficiency, makes it suitable for applying a consistent theoretical treatment across related peptide structures [14–16]. The energetic consequences of peptide association can be examined by comparing the electronic energies of independently optimised monomeric and dimeric states. The resulting dimerization energy incorporates the energetic consequences of intermolecular association together with structural reorganisation accompanying formation of the optimised dimer. This distinction is particularly relevant to flexible peptides, for which association can be accompanied by substantial conformational rearrangement. Sequence dependent comparison of these energies therefore provides a quantitative means of examining how residue arrangement influences the energetic outcome of molecular association within consistently defined molecular models. Establishing the contribution of residue order requires sequence arrangement to be distinguished from amino acid composition. Composition matched scrambled controls provide a direct basis for this comparison because the constituent amino acids are conserved while their linear arrangement is altered. Differences between an aggregation prone sequence and its composition matched scrambled control can therefore reveal how residue order reorganises the molecular interactions accessible to the same amino acid composition. Within a common computational framework, this comparison enables sequence encoded effects to be examined through molecular geometry, structural reorganisation and dimerization energetics.

Energetic differences provide one dimension of molecular association, while the spatial distribution of electronic properties provides another. Molecular electrostatic potential (ESP) describes the electrostatic potential generated by the molecular charge distribution and identifies spatial regions of positive and negative potential [17, 18]. Mapping ESP onto a defined electron density surface enables the electrostatic environments presented by related molecular structures to be compared directly [17, 18]. These electrostatic patterns are relevant to molecular recognition because the position and magnitude of local potential regions influence the orientation and complementarity of interacting molecules. Comparison of aggregation prone sequences with composition matched controls can therefore establish whether residue rearrangement produces distinct electrostatic landscapes that accompany differences in association energetics.

Density functional theory (DFT) provides an electronic structure framework for calculating ground state electron density and properties derived from it. The Grid based Projector Augmented Wave (GPAW) method implements DFT using the projector augmented wave formalism and supports real space grid representations suitable for analysing molecular electron density and electrostatic properties [19, 20]. Applying identical electronic structure settings across aggregation prone peptides and their composition matched controls provides internally consistent electron density and ESP data for comparative analysis. These spatial electronic properties can be complemented by Hirshfeld population analysis, which partitions molecular electron density among individual atoms through stockholder weighting relative to reference atomic densities [21]. The resulting atomic charges provide an atom resolved representation of molecular charge distribution, allowing differences in charge magnitude, dispersion and heterogeneity to be quantified between related peptide structures [21, 22]. Together, dimerization energetics, electron density, surface ESP and Hirshfeld charge distributions provide complementary descriptions of sequence dependent molecular behaviour. Dimerization energy quantifies the energetic difference between independently optimised monomeric and associated states, while electron density describes the spatial distribution of electrons within the molecular structure. Surface ESP describes the electrostatic environment generated by that distribution, and Hirshfeld analysis resolves the distribution into atom specific charge contributions. Integrating these properties provides a coherent framework for determining whether changes in residue order produce corresponding changes in molecular structure, association energetics and electronic organisation.

In this study, aggregation prone peptide segments derived from α -synuclein and A β were investigated alongside composition matched scrambled controls to determine how residue order influences early molecular association. Monomer and homodimer structures were optimised using GFN2-xTB, and dimerization energies were calculated from the corresponding optimised electronic energies. DFT calculations using GPAW were subsequently applied to the optimised monomers to determine electron density and surface ESP, while Hirshfeld population analysis was used to quantify atom resolved charge distributions. By integrating structural optimisation, dimerization energetics and spatial and atom resolved electronic properties within matched sequence comparisons, the study establishes a quantitative framework for identifying sequence encoded molecular determinants of early peptide association in α -synuclein and A β .

Computational methods

Atomistic models of the control and aggregation prone α -synuclein and A β peptides were investigated using a consistent computational workflow comprising sequence analysis, molecular geometry optimization, dimer association calculations, structural and interface analysis, DFT, Hirshfeld charge analysis and ESP analysis. Open Babel was used for molecular structure conversion during structure preparation [23], and the structures were checked for atomic connectivity, residue composition and terminal capping before geometry optimisation. Calculations and data processing were performed using Python 3.11.14, xTB 6.7.1, GPAW 24.6.0, Atomic Simulation Environment (ASE) 3.26.0,[24] NumPy 1.26.4 [25], SciPy 1.17.0, pandas [26], Matplotlib [27] and scikit image. GPAW employed LibXC 6.2.2 and the corresponding projector augmented wave (PAW) datasets. The peptide structures were geometry optimized using the GFN2-xTB method. For intermolecular association calculations, two copies of each optimized monomer were used to construct the corresponding dimer with an initial intermolecular separation of 3.5 Å. The dimer structures were initially relaxed using the GFN force field, followed by geometry optimization using GFN2-xTB. The molecular charge and optimization criteria appropriate to each peptide system were maintained throughout the calculations. Dimer association energies, ΔE_{assoc} , were calculated from the converged GFN2-xTB energies according to equation (Eq.) 1:

$$\Delta E_{assoc} = E_{dimer} - 2E_{monomer} \quad (1)$$

where E_{dimer} is the total energy of the optimized dimer and $E_{monomer}$ is the total energy of the corresponding optimized monomer. Energies obtained in Hartree were converted to electronvolts for comparative analysis. Sequence descriptors were calculated directly from the amino acid sequences of the control and aggregation prone peptides. Residue charge assignments were used to determine net sequence charge and the positional distribution of charged residues. Residue specific hydropathy values were used to characterize hydropathy distribution, local hydrophobic clustering and hydrophobic run length. Sequence charge decoration was calculated from residue charge identities and their sequence separation to quantify the spatial patterning of charge along each peptide sequence. Identical definitions were applied to the aggregation prone peptide and its corresponding composition matched scrambled control. Structural descriptors were calculated from the optimized monomer coordinates. The radius of gyration (R_g) was determined relative to the molecular centre of mass as a measure of molecular compactness. The maximum molecular dimension and end-to-end distance were calculated from the optimized coordinates. Molecular shape was characterized from the principal dimensions of the gyration tensor, and asphericity was calculated to quantify deviation from a spherical atomic distribution. A compactness index was obtained by normalizing R_g to the cube root of the number of atoms.

Intermolecular interfaces were characterized from the optimized dimer geometries using pairwise distances between atoms belonging to opposing monomers. The minimum intermolecular distance was determined for each dimer. Total intermolecular contacts were evaluated at ≤ 4.0 Å, while heavy atom contacts were quantified at ≤ 4.5 Å and ≤ 5.0 Å. Close intermolecular contacts involving N, O or S atoms were identified at ≤ 3.5 Å. Heavy contact density was calculated from the corresponding intermolecular

contact population to provide a normalized measure of interface packing. Electronic structure calculations were performed on the optimised monomer structures using DFT as implemented in GPAW, with ASE used for atomic structure handling and calculator interfacing. The Perdew Burke Ernzerhof (PBE) exchange correlation functional was employed within the PAW formalism. Calculations used a finite difference representation with a real space grid spacing of 0.20 Å. Each peptide was centred within a simulation cell containing 6.0 Å of vacuum around the molecular structure. Molecular charges of -2 and +1 were assigned to the Aβ and α-synuclein systems, respectively. Spin polarization was assigned according to the electron count of each charged system. Electronic occupations were treated using Fermi Dirac smearing of 0.05 eV, and self-consistent electronic calculations were converged using a density criterion of 1.0×10^{-4} . Hirshfeld population analysis was performed from the converged electronic densities to quantify atomic charge distributions. Atomic Hirshfeld charges were extracted for each peptide and assigned to the corresponding atomic coordinates and elemental identities. The distributions were characterized using the total, mean, median, minimum and maximum charge, standard deviation, mean absolute charge, positive and negative atom fractions, positive and negative charge sums and percentile statistics. Residue resolved charge distributions were obtained by summing the atomic Hirshfeld charges associated with each residue. Three dimensional all electron density and ESP grids were obtained from the converged GPAW calculations. The molecular surface was defined using an electron density isovalue of 0.001 e Bohr⁻³, and the corresponding ESP values were mapped onto the density surface. Surface ESP distributions were quantified using the minimum, maximum, mean, median, standard deviation, root mean square, mean absolute value and percentile statistics. A common electrostatic scale was used for comparative visualization of the four peptide systems.

Relative ESP was calculated independently for each peptide by mean centring the surface ESP distribution according to Eq. 2:

$$ESP_{\{relative,i\}} = ESP_i - \overline{ESP}$$

2

where ESP_i is the ESP at surface vertex i and $\overline{ESP}$ is the mean ESP across all surface vertices of the corresponding peptide. Relative electrostatic heterogeneity was characterised from the standard deviation, root mean square, mean absolute value, range and percentile distribution of the mean centred ESP values. Positive and negative surface fractions were determined from the sign of the relative ESP, and strongly positive and strongly negative surface regions were quantified using thresholds of $\geq +1$ eV and ≤ -1 eV, respectively. All comparative descriptors were calculated using identical definitions and numerical criteria for each aggregation prone peptide and its corresponding scrambled control.

Results and Discussion

Geometry optimization and monomer energetics

Geometry optimization was performed to obtain consistently optimized monomer structures for the aggregation prone peptides and their composition matched scrambled controls, providing the structural and energetic reference states for subsequent comparative analyses. For α -synuclein, the aggregation prone sequence VTGVTAVAQK was compared with the composition matched scrambled sequence AGATQTVVVK, while the A β aggregation prone sequence LVFFAEDVGS was compared with EFLVGSDVAF. Within each peptide pair, peptide length and amino acid composition were conserved while residue order was altered, allowing sequence dependent differences in optimized geometry and monomer energetics to be evaluated under identical computational conditions. The optimized structures comprised A β scrambled control (Fig. 1A), A β aggregation prone peptide (Fig. 1B), α -synuclein scrambled control (Fig. 1C), and α -synuclein aggregation prone peptide (Fig. 1D). The A β scrambled control and aggregation prone peptide each contained 158 atoms and carried a molecular charge of -2 (Table 1). Following optimization, the scrambled control exhibited a GFN2-xTB electronic energy of -252.8816448 Eh, compared with -252.8643808 Eh for the aggregation prone peptide. The energy difference between the two optimized monomers was 0.0172640 Eh, equivalent to approximately 45.33 kJ/mol, with the scrambled control exhibiting the lower electronic energy. Since amino acid composition, atom number and molecular charge were conserved within the A β pair, the difference was associated with the distinct intramolecular configurations arising from residue rearrangement. The corresponding conformational differences were evident in the optimized structures shown in Fig. 1A and B. The α -synuclein scrambled control and aggregation prone peptide each contained 154 atoms and carried a molecular charge of +1 (Table 1). The optimized scrambled control exhibited a GFN2-xTB electronic energy of -232.9472350 Eh, whereas the aggregation prone peptide exhibited an energy of -233.0047931 Eh. The aggregation prone peptide was therefore lower in electronic energy by 0.0575580 Eh, equivalent to approximately 151.12 kJ/mol. This energetic ordering contrasted with the A β pair, for which the scrambled control exhibited the lower electronic energy. Thus, rearrangement of an identical amino acid composition produced different energetic responses in the two peptide systems, reflecting sequence dependent differences in the intramolecular configurations sampled during optimization. The optimization characteristics of the four monomers are summarized in Table 1. The A β scrambled control and aggregation prone peptide exhibited gradient norms of 0.000787571 and 0.000573192 Eh a_0^{-1} after 224 and 201 optimization iterations, respectively. The corresponding values for the α -synuclein scrambled control and aggregation prone peptide were 0.000705170 and 0.000450745 Eh a_0^{-1} after 264 and 380 iterations, respectively. These optimized geometries and their associated electronic energies were subsequently used for structural characterization and as the monomer reference states for calculation of dimer association energies. In general, residue rearrangement produced distinct optimized conformations and monomer electronic energies despite conservation of amino acid composition within each matched pair. The A β scrambled control exhibited an electronic energy approximately 45.33 kJ/mol lower than the aggregation prone peptide, whereas the α -synuclein aggregation prone peptide exhibited an electronic energy approximately 151.12 kJ/mol lower than its scrambled control (Table 1). The contrasting energetic ordering showed that the effect of residue sequence on the optimized monomer state differed between A β and α -synuclein, providing the structural and energetic reference for subsequent analysis of their intermolecular association.

Table 1

Final electronic energies and geometry optimization parameters for the A β and α -synuclein aggregation prone peptides and their composition matched scrambled controls calculated using GFN2-xTB.

Peptide	Charge	GFN2 energy (Eh)	Gradient norm (Eh a ₀ ⁻¹)	Optimization iterations	Number of atoms
Amyloid β scrambled control	-2	-252.8816448	0.000787571	224	158
Amyloid β aggregation prone peptide	-2	-252.8643808	0.000573192	201	158
α -Synuclein scrambled control	+ 1	-232.947235	0.00070517	264	154
α -Synuclein aggregation prone peptide	+ 1	-233.0047931	0.000450745	380	154

Dimer structures and association energetics

Dimerization was examined to determine how residue order influenced intermolecular association when amino acid composition was conserved within each peptide pair. The optimized monomers provided the corresponding energetic reference states for calculation of the association energies, while structural descriptors derived from the optimized dimers were used to characterize the intermolecular interfaces. The association energies and principal interface characteristics of the A β and α -synuclein dimers are summarized in Table 2. For A β , the scrambled control and aggregation prone peptide exhibited association energies of 232.552 and 115.617 kJ mol⁻¹, respectively (Table 2). The aggregation prone sequence therefore showed an association energy that was 116.935 kJ mol⁻¹ lower than that of the composition matched scrambled control. Both values were positive relative to their corresponding optimized monomer reference states, indicating that the optimized dimers were higher in electronic energy than the two independently optimized monomers. The substantially lower value obtained for the aggregation prone sequence showed that residue rearrangement altered the energetic outcome of association within the A β pair. The structural characteristics of the A β interfaces showed a different trend. The minimum heavy atom intermonomer distance was 3.105 Å for the scrambled control and 3.192 Å for the aggregation prone peptide. The scrambled control contained 81 heavy atom contacts within 4.5 Å, compared with 40 for the aggregation prone peptide, while both dimers contained one intermolecular N, O or S contact within 3.5 Å. The corresponding heavy contact densities were 1.000 and 0.494, respectively. Thus, the lower association energy of the A β aggregation prone sequence was accompanied by fewer close heavy atom contacts and a lower contact density. The calculated association energy therefore did not vary directly with the number of geometrically defined intermolecular contacts within the A β pair. A different relationship between association energetics and interface organization was observed for α -synuclein. The scrambled control exhibited an association

energy of -494.371 kJ/mol, whereas the aggregation prone peptide exhibited an association energy of -172.819 kJ/mol (Table 2). Both values were negative relative to their corresponding optimized monomer reference states. The scrambled control exhibited the more favourable association energy, with a difference of 321.552 kJ/mol between the two sequences. Residue rearrangement therefore produced a substantial change in the energetic outcome of α -synuclein dimerization. In contrast to the energetic ordering, the α -synuclein aggregation prone peptide exhibited the more extensive intermolecular interface. Its minimum heavy atom intermonomer distance was 2.705 Å, compared with 2.761 Å for the scrambled control. Heavy atom contacts within 4.5 Å increased from 105 in the scrambled control to 166 in the aggregation prone peptide, while intermolecular N, O or S contacts within 3.5 Å increased from 5 to 15. Heavy contact density similarly increased from 1.458 to 2.306. Residue ordering in the aggregation prone α -synuclein peptide therefore produced closer intermolecular approach and greater interface contact density, although its association energy was less favourable than that of the scrambled control. Comparison of the two peptide families demonstrated that residue order influenced association through distinct structural and energetic responses. For A β , the aggregation prone sequence exhibited a substantially lower positive association energy than the scrambled control despite having fewer heavy atom contacts and a lower contact density. For α -synuclein, the aggregation prone sequence exhibited closer intermolecular approach, more heavy atom contacts, more close N, O or S contacts and greater contact density, whereas the scrambled control exhibited the more negative association energy. Association energy and interface packing therefore represented complementary characteristics of dimer formation rather than equivalent measures of association. These differences are relevant to the energetic definition used in this study. Association energy represented the electronic energy difference between the optimized dimer and its independently optimized monomer reference states and consequently incorporated the energetic consequences associated with formation of the optimized dimer. In contrast, the interface descriptors quantified geometrical characteristics of intermolecular organization within the optimized dimer. The different relationships between these parameters in A β and α -synuclein showed that the extent of close intermolecular contact alone did not account for the calculated energetic outcome of dimerization. In general, the composition matched comparisons demonstrated that changing residue order altered both dimer association energetics and intermolecular interface organization. The A β aggregation prone sequence was distinguished by a lower association energy relative to its scrambled control, whereas the α -synuclein aggregation prone sequence was distinguished by closer intermolecular approach and a more extensively populated interface. These findings showed that sequence order influenced early molecular association through peptide specific changes in both energetic and structural organization, providing direct evidence that the arrangement of residues, rather than amino acid composition alone, contributed to the molecular determinants of association in the A β and α -synuclein model systems.

Table 2

Association energetics and intermolecular interface characteristics of the optimized peptide dimers.

Peptide	Dimer charge	Association energy (kJ/mol)	Minimum heavy atom intermonomer distance (Å)	Heavy atom contacts ≤ 4.5 Å	Intermonomer N/O/S contacts ≤ 3.5 Å	Heavy contact density
Amyloid β scrambled control	-4	232.552	3.105	81	1	1.000
Amyloid β aggregation prone peptide	-4	115.617	3.192	40	1	0.494
α -Synuclein scrambled control	+ 2	-494.371	2.761	105	5	1.458
α -Synuclein aggregation prone peptide	+ 2	-172.819	2.705	166	15	2.306

Electrostatic potential landscapes

ESP analysis was used to examine how residue order influenced the spatial electrostatic characteristics of the composition matched A β and α -synuclein peptides. The relative ESP maps corresponded to the A β scrambled control (Fig. 2A), A β aggregation prone peptide (Fig. 2B), α -synuclein scrambled control (Fig. 2C), and α -synuclein aggregation prone peptide (Fig. 2D). Because amino acid composition and molecular charge were conserved within each aggregation prone and scrambled control pair, differences in the ESP landscapes reflected changes arising from residue order and the resulting 3D molecular organization. The ESP values were mean centred independently for each peptide before visualization. Accordingly, red regions represented surface ESP values above the mean of the corresponding peptide, whereas blue regions represented values below its mean. The common relative ESP scale from -3 to $+3$ eV therefore provided a consistent representation of spatial ESP variation across the four structures. The mapped values represented deviations from the mean ESP of each individual peptide and were consequently interpreted in terms of the spatial distribution and heterogeneity of relative ESP within each molecular surface. The A β scrambled control exhibited a heterogeneous ESP landscape, with positive and negative relative ESP distributed across different regions of the molecular surface (Fig. 2A). A prominent region of positive relative ESP was concentrated toward one end of the structure, while negative relative ESP occurred across several other surface regions. Rearrangement of the same amino acid composition in the A β aggregation prone peptide produced a distinctly different ESP landscape (Fig. 2B). Extended regions of negative relative ESP were apparent across one part of the structure, while positive relative ESP occupied other spatially separated regions. The transition from Fig. 2A to Fig. 2B therefore demonstrated that residue order altered the location, extent and spatial organization of relative

ESP despite conservation of amino acid composition and molecular charge. The sequence dependent redistribution of ESP in A β was accompanied by a substantial difference in association energetics. The aggregation prone peptide exhibited an association energy of 115.617 KJ/mol compared with 232.552 kJ/mol for the scrambled control, corresponding to a decrease of 116.935 KJ/mol. Thus, the altered ESP landscape of the aggregation prone sequence occurred alongside a lower calculated association energy. These results showed that rearrangement of residue order affected both the spatial organization of ESP and the energetics of dimer association. The relationship was interpreted collectively rather than as a direct correspondence between individual positive or negative ESP regions and the calculated association energy, since association reflected the combined molecular interactions established within the dimer.

A distinct sequence dependent ESP response was also observed for α -synuclein. The scrambled control displayed broad spatial separation of regions with positive and negative relative ESP (Fig. 2C). Positive relative ESP extended across a substantial region of the structure, while increasingly negative relative ESP was localized toward the opposing part of the molecular surface. The α -synuclein aggregation prone peptide exhibited a different spatial distribution of these regions (Fig. 2D). Negative relative ESP was pronounced toward one terminal region, whereas positive relative ESP was concentrated predominantly across the opposing portion of the structure. Comparison of Fig. 2C and Fig. 2D therefore showed that residue rearrangement modified the position, continuity and spatial extent of the ESP regions presented by the optimized α -synuclein structures.

The surface ESP probability density distributions further quantified these sequence dependent electrostatic differences (Fig. 2E and F). For A β , both sequences were dominated by positive ESP values, but their distributions differed markedly (Fig. 2E). The scrambled control showed a dominant population within the positive ESP region around 2 to 3 eV, whereas the aggregation prone peptide exhibited a prominent population at lower positive ESP and a broader distribution extending toward higher positive potentials. Although the profiles overlapped across much of the positive ESP range, differences in their probability densities demonstrated redistribution of the surface electrostatic environments following residue rearrangement. For α -synuclein, both sequences were dominated by negative ESP values and exhibited broad, multimodal distributions (Fig. 2F). The aggregation prone peptide showed greater probability density within part of the more negative ESP region, particularly around -3 eV, whereas the scrambled control showed greater probability density across portions of the intermediate negative region. Considerable overlap remained between the distributions, but their differing probability densities demonstrated sequence dependent redistribution of surface ESP. Collectively, the spatial maps and probability density distributions provided complementary descriptions of the electrostatic consequences of residue rearrangement. Figure 2A - D resolved the spatial organization of relative ESP across the optimized molecular surfaces, while Fig. 2E and F quantified the distributions of surface ESP values. The differences observed within both composition matched pairs established that amino acid composition and molecular charge alone did not determine the resulting electrostatic landscape. Rather, residue order altered the 3D arrangement of the constituent chemical groups and consequently changed both the spatial organization and distribution of surface electrostatic potential. These findings identified

sequence dependent electrostatic organization as a molecular property through which peptides of identical composition acquired distinct physicochemical characteristics relevant to their association behaviour.

Hirshfeld atomic charge distribution and charge polarisation

Hirshfeld atomic charge analysis revealed sequence dependent differences in the distribution of electronic charge across the α -synuclein and A β systems. The probability distributions showed that both aggregation prone peptides and their scrambled controls contained populations of positively and negatively charged atoms, with the greatest probability concentrated within the positive charge region (Fig. 3A and B). For α -synuclein, the aggregation prone and scrambled sequences exhibited broadly similar charge ranges but differed in the relative probabilities across these ranges (Fig. 3A). The scrambled control showed a higher probability within the principal positive charge region near 0.04 e, whereas the aggregation prone peptide displayed a lower probability in this region together with redistribution across adjacent positive and negative charge intervals. A β showed a related sequence dependent response, with the scrambled control exhibiting a pronounced probability maximum near 0.04 e and the aggregation prone peptide showing a reduced maximum accompanied by redistribution across the wider charge range (Fig. 3B). These changes demonstrated that sequence rearrangement modified the atomic charge distribution in both peptide systems. The individual Hirshfeld charges further resolved these differences at the atomic level (Fig. 3C). All four systems contained distinct populations of positive and negative atomic charges, with the aggregation prone and scrambled sequences occupying comparable overall charge domains but differing in their distribution within them. The A β systems extended to more negative charges than the corresponding α -synuclein systems, consistent with a broader negative charge domain within A β . Importantly, the sequence dependent differences observed in Fig. 3A and B arose from redistribution of atomic charge populations rather than from a uniform displacement of all atomic charges in one direction. This provided an atom resolved description of how sequence order altered the electronic environment of the peptides. The balance between positive and negative atomic populations provided an additional quantitative distinction between the systems (Fig. 3D). Positive atoms constituted approximately 69% of both α -synuclein sequences, while negative atoms accounted for approximately 31%, indicating that sequence rearrangement produced little change in the overall proportion of positive and negative atoms in this system. In contrast, the A β aggregation prone peptide contained approximately 62% positive and 38% negative atoms, compared with approximately 64% positive and 36% negative atoms for the scrambled control. Thus, the A β sequence showed a modest change in the balance of positive and negative atomic populations, whereas the corresponding α -synuclein fractions remained closely conserved. Charge magnitude and dispersion further distinguished changes in atomic charge organisation from changes in the overall extent of charge polarisation (Fig. 3E). For α -synuclein, the aggregation prone peptide exhibited a mean absolute charge of approximately 0.084 e and a charge standard deviation of 0.106 e, compared with approximately 0.086 e and 0.109 e, respectively, for the scrambled control. The corresponding values for the A β aggregation prone peptide were approximately 0.082 e and 0.112 e, compared with approximately

0.082 e and 0.113 e for the scrambled control. The close mean absolute charges and charge dispersions within each composition matched pair showed that residue rearrangement produced comparatively small changes in global charge magnitude and overall charge variability. When considered alongside the distinct probability distributions in Fig. 3A and B, these results showed that the principal electronic response to residue rearrangement involved redistribution of atomic charge populations rather than substantial alteration of their overall magnitude. Collectively, the Hirshfeld analysis established atomic charge organisation as a sequence sensitive electronic property. In α -synuclein, similar positive and negative atom fractions, charge magnitudes and charge dispersions were accompanied by distinct atomic charge probability distributions, demonstrating electronic reorganisation despite conservation of these global charge characteristics. A β similarly exhibited redistribution of atomic charge, together with a modest change in the balance of positive and negative atomic populations. Combined with the sequence dependent ESP landscapes, these findings showed that residue order reorganised the atom resolved and spatial presentation of electrostatic character while largely conserving global atomic charge magnitude. This electronic reorganisation provides a molecular basis through which identical amino acid compositions can generate distinct electrostatic environments and complements the sequence dependent differences in structure, intermolecular organisation and association energetics observed for the α -synuclein and A β systems.

Sequence patterning, hydrophobic organisation and molecular shape

Sequence based descriptors showed that residue rearrangement altered the distribution of charge and hydrophobic character within the composition matched peptide pairs. Sequence charge decoration (SCD) was zero for both α -synuclein sequences, whereas A β exhibited values of 0.100 for the aggregation prone sequence and 0.245 for the scrambled control (Fig. 4A). The difference within the A β pair demonstrated that rearrangement of an identical amino acid composition changed the sequence patterning of charged residues, while the corresponding α -synuclein rearrangement produced no change in this descriptor. The effect of residue order on charge patterning was therefore peptide dependent. Local hydrophobic organisation showed a distinct pattern (Fig. 4B). The maximum three residue mean hydropathy increased from approximately 2.6 in aggregation prone α -synuclein to approximately 4.2 in the scrambled control. In contrast, the aggregation prone and scrambled A β sequences exhibited similar maximum local hydropathy values of approximately 3.6. Hydrophobic run length provided an additional distinction (Fig. 4C). Both α -synuclein sequences contained a maximum hydrophobic run of three residues, whereas aggregation prone A β contained a five-residue hydrophobic run compared with three residues in the scrambled control. Thus, the A β aggregation prone sequence maintained hydrophobic character across a longer contiguous sequence region despite having a similar maximum local hydropathy to its composition matched control. These results showed that residue order influenced different aspects of hydrophobic organisation in the two peptide systems. Structural descriptors were subsequently examined to determine whether the sequence differences were accompanied by changes in global molecular dimensions and shape. The radius of gyration (R_g), was approximately 11.1 Å for

aggregation prone α -synuclein and 10.9 Å for the scrambled control (Fig. 4D), indicating closely comparable overall dimensions. A larger difference was observed for A β , where the aggregation prone peptide exhibited an Rg of approximately 10.1 Å compared with approximately 11.0 Å for the scrambled control, consistent with a more compact optimized structure. End-to-end distances remained closely similar across the four systems, at approximately 36 to 37 Å (Fig. 4E). Molecular asphericity also varied only modestly within each composition matched pair, with values of approximately 0.89 and 0.88 for aggregation prone and scrambled α -synuclein, respectively, and approximately 0.84 and 0.85 for aggregation prone and scrambled A β , respectively (Fig. 4F). Collectively, the sequence and structural descriptors showed that residue rearrangement produced more pronounced changes in local physicochemical patterning than in overall molecular dimensions. For α -synuclein, sequence scrambling substantially changed the maximum local hydrophathy while SCD, Rg, end-to-end distance and asphericity remained closely comparable. For A β , residue rearrangement altered SCD, hydrophobic run length and Rg, while maximum local hydrophathy, end-to-end distance and asphericity remained comparatively similar. The A β aggregation prone sequence was therefore distinguished from its scrambled control by a different charge patterning metric, a longer contiguous hydrophobic segment and a smaller Rg. The sequence dependent changes in physicochemical patterning differed between the two peptide systems. In α -synuclein, residue rearrangement predominantly altered local hydrophathy, whereas in A β it altered charge patterning, hydrophobic run length and molecular compactness. These differences, together with the association energetics, ESP landscapes and Hirshfeld charge distributions, showed that residue order influenced multiple molecular properties associated with early peptide association despite conservation of amino acid composition. The distinct responses of α -synuclein and A β further indicated that the molecular consequences of residue rearrangement depended on the physicochemical organisation of the individual sequence.

Intermolecular contact organisation at the peptide association interface

Intermolecular contact analysis showed that residue order altered the organisation of the association interface in both peptide systems. For α -synuclein, the aggregation prone dimer contained 166 heavy atom contacts within 4.5 Å, compared with 105 for the scrambled control (Fig. 5A). The number of intermolecular contacts involving N, O or S atoms within 3.5 Å increased from 5 in the scrambled control to 15 in the aggregation prone dimer (Fig. 5B). Thus, the aggregation prone α -synuclein sequence produced a more extensively populated interface, with increases in both total heavy atom contacts and close N, O or S contacts. A β exhibited the opposite trend in total interface contact number. The aggregation prone dimer contained 40 heavy atom contacts within 4.5 Å, compared with 81 for the scrambled control (Fig. 5A). Both A β dimers contained one intermolecular N, O or S contact within 3.5 Å (Fig. 5B). Residue rearrangement therefore altered the extent of the A β intermolecular interface without changing the number of close N, O or S contacts detected at the applied distance criterion. The minimum heavy atom intermonomer distance further differentiated the two peptide systems (Fig. 5C). The aggregation prone and scrambled α -synuclein dimers exhibited minimum distances of 2.705 and

2.761 Å, respectively, whereas the corresponding Aβ values were 3.192 and 3.105 Å. The shorter minimum distances observed for α-synuclein were consistent with the greater number of close intermolecular contacts identified in this peptide system. Within each composition matched pair, however, the difference in minimum distance was modest relative to the change in heavy atom contact number, indicating that interface organisation was more clearly distinguished by the number and type of contacts than by minimum distance alone. In summary, the results showed that residue order reorganised intermolecular contacts in a peptide dependent manner despite conservation of amino acid composition. Aggregation prone α-synuclein formed a substantially more populated interface than its scrambled control, whereas aggregation prone Aβ formed fewer heavy atom contacts than its control while retaining the same number of close N, O or S contacts. These contrasting responses demonstrated that the structural consequences of residue rearrangement differed between α-synuclein and Aβ and complemented the distinct association energetics observed for the same composition matched pairs. The interface descriptors therefore provided a structural basis for relating sequence order to the early intermolecular organisation of the two peptide systems.

Schematic representation of the computational framework

Scheme 1 summarised the integrated computational framework used to examine sequence dependent early peptide association in α-synuclein and Aβ. The analysis began with aggregation prone peptide sequences and their composition matched scrambled controls, allowing residue order to be varied while amino acid composition was conserved. The corresponding monomer structures were then optimised using GFN2-xTB, providing the reference geometries and electronic energies for the subsequent comparative analyses. Homodimers were constructed from the optimised monomers and further optimised to determine association energetics and intermolecular organisation. Electronic structure analysis was performed on the optimised monomers using DFT, from which ESP distributions and Hirshfeld atomic charges were obtained. These calculations resolved sequence dependent changes in the spatial organisation of electrostatic potential and in atom resolved charge distribution. The molecular and interfacial analyses were then integrated through association energies, intermolecular contact metrics, Hirshfeld charge distributions and sequence descriptors, including charge decoration, local hydropathy, hydrophobic run length and radius of gyration. The scheme therefore connected residue order with successive levels of molecular organisation, from sequence patterning and monomer structure to dimer association, intermolecular contacts and electronic structure. This integrated comparison showed that rearrangement of identical amino acid compositions produced peptide specific changes in energetics, intermolecular organisation and electrostatic character, establishing residue order as a determinant of the early association behaviour of α-synuclein and Aβ.

Conclusions

This study showed that residue order influenced the structural, energetic and electronic properties associated with early peptide association despite conservation of amino acid composition. Aggregation

prone A β and α -synuclein exhibited distinct responses to sequence rearrangement, with differences in association energetics, intermolecular contact organisation, ESP landscapes, Hirshfeld charge distributions and sequence dependent physicochemical properties. Notably, greater interface contact abundance did not correspond to a common trend in association energy across the two peptide systems, indicating that these descriptors captured complementary aspects of molecular association. The combined analyses further showed that residue rearrangement reorganised electrostatic character while global atomic charge magnitude remained comparatively conserved within each peptide pair. Collectively, these findings demonstrated that residue order shaped early association through peptide specific changes in molecular structure, electrostatic organisation and intermolecular interactions, providing a molecular basis for the distinct association behaviour of α -synuclein and A β .

Acknowledgments OA gratefully acknowledges funding from the Engineering and Physical Sciences Research Council (EPSRC), grant EP/X029956/1. O.J.A. gratefully acknowledges financial support from the Royal Society through grant numbers RG\R2\232090 & ICAO/R1/231012 (Research Grant & International Science Partnership Fund (ISPF) – International Collaboration Awards).

Declarations

Acknowledgments OA gratefully acknowledges funding from the Engineering and Physical Sciences Research Council (EPSRC), grant EP/X029956/1. O.J.A. gratefully acknowledges financial support from the Royal Society through grant numbers RG\R2\232090 & ICAO/R1/231012 (Research Grant & International Science Partnership Fund (ISPF) – International Collaboration Awards).

Data availability All data will be made available upon request.

Competing interests The authors declare no competing interests.

Author Contribution

OA: Conceptualization, methodology development, model construction, computational implementation, data curation, formal analysis, visualization, software scripting, interpretation of results, and original draft preparation; OJA: critical review, scientific editing, intellectual input, and validation of the final manuscript

References

1. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–366
2. Knowles TPJ, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* 15:384–396
3. Eisenberg D, Jucker M (2012) The amyloid state of proteins in human diseases. *Cell* 148:1188–1203

4. Soto C, Pritzkow S (2018) Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* 21:1332–1340
5. Cline EN, Bicca MA, Viola KL, Klein WL (2018) The Amyloid- β Oligomer Hypothesis: Beginning of the Third Decade. *Journal of Alzheimer's Disease* 64:S567–S610
6. Uversky VN (2015) Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front Aging Neurosci* 7:. <https://doi.org/10.3389/fnagi.2015.00018>
7. Rodriguez JA, Ivanova MI, Sawaya MR, et al (2015) Structure of the toxic core of α -synuclein from invisible crystals. *Nature* 525:486–490. <https://doi.org/10.1038/nature15368>
8. Colvin MT, Silvers R, Ni QZ, et al (2016) Atomic Resolution Structure of Monomorphic A β 42 Amyloid Fibrils. *J Am Chem Soc* 138:9663–9674. <https://doi.org/10.1021/jacs.6b05129>
9. Knowles TPJ, Waudby CA, Devlin GL, et al (2009) An analytical solution to the kinetics of breakable filament assembly. *Science* (1979) 326:1533–1537. <https://doi.org/10.1126/science.1178250>
10. Dear AJ, Meisl G, Šarić A, et al (2020) Identification of on- And off-pathway oligomers in amyloid fibril formation. *Chem Sci* 11:6236–6247. <https://doi.org/10.1039/c9sc06501f>
11. Cawood EE, Karamanos TK, Wilson AJ, Radford SE (2021) Visualizing and trapping transient oligomers in amyloid assembly pathways. *Biophys Chem* 268:. <https://doi.org/10.1016/j.bpc.2020.106505>
12. Smaoui MR, Poitevin F, Delarue M, et al (2013) Computational assembly of polymorphic amyloid fibrils reveals stable aggregates. *Biophys J* 104:683–693. <https://doi.org/10.1016/j.bpj.2012.12.037>
13. Robustelli P, Piana S, Shaw DE (2018) Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* 115:E4758–E4766. <https://doi.org/10.1073/pnas.1800690115>
14. Bannwarth C, Ehlert S, Grimme S (2019) GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J Chem Theory Comput* 15:1652–1671. <https://doi.org/10.1021/acs.jctc.8b01176>
15. Bannwarth C, Caldeweyher E, Ehlert S, et al (2021) Extended tight-binding quantum chemistry methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 11
16. Grimme S, Bannwarth C, Shushkov P (2017) A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements ($Z = 1-86$). *J Chem Theory Comput* 13:1989–2009. <https://doi.org/10.1021/acs.jctc.7b00118>
17. Politzer P, Murray JS (2002) The fundamental nature and role of the electrostatic potential in atoms and molecules. *Theor. Chem. Acc.* 108:134–142
18. Murray JS, Politzer P (2011) The electrostatic potential: An overview. *Wiley Interdiscip Rev Comput Mol Sci* 1:153–163. <https://doi.org/10.1002/wcms.19>

19. Enkovaara J, Rostgaard C, Mortensen JJ, et al (2010) Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method. *Journal of Physics Condensed Matter* 22
20. Mortensen JJ, Larsen AH, Kuisma M, et al (2024) GPAW: An open Python package for electronic structure calculations. *Journal of Chemical Physics* 160:. <https://doi.org/10.1063/5.0182685>
21. Hirshfeld FL (1977) XVII. Spatial Partitioning of Charge Density. *Isr J Chem* 16:198–201. <https://doi.org/10.1002/ijch.197700033>
22. Manz TA, Limas NG (2016) Introducing DDEC6 atomic population analysis: Part 1. Charge partitioning theory and methodology. *RSC Adv* 6:47771–47801. <https://doi.org/10.1039/c6ra04656h>
23. O’Boyle NM, Banck M, James CA, et al (2011) Open Babel: An Open chemical toolbox. *J Cheminform* 3:. <https://doi.org/10.1186/1758-2946-3-33>
24. Hjorth Larsen A, Jørgen Mortensen J, Blomqvist J, et al (2017) The atomic simulation environment - A Python library for working with atoms. *Journal of Physics Condensed Matter* 29
25. Harris CR, Millman KJ, van der Walt SJ, et al (2020) Array programming with NumPy. *Nature* 585:357–362
26. McKinney W (2010) Data Structures for Statistical Computing in Python. pp 56–61
27. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9:90–95. <https://doi.org/10.1109/MCSE.2007.55>

Scheme 1

Scheme 1 is available in the supplementary files section

Figures

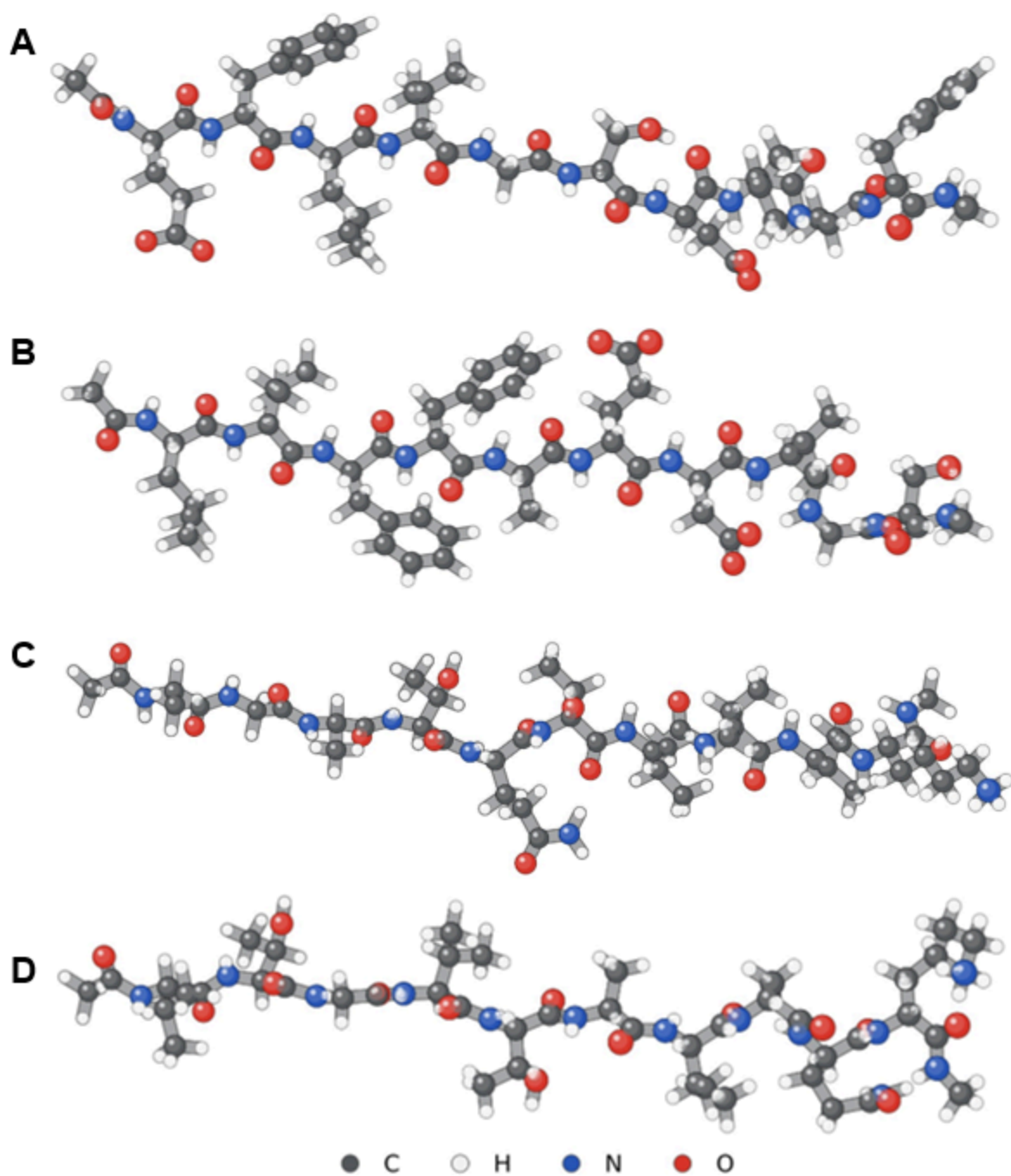


Figure 1

Optimized monomer structures of (A) A β scrambled control, (B) A β aggregation prone peptide, (C) α -synuclein scrambled control, and (D) α -synuclein aggregation prone peptide.

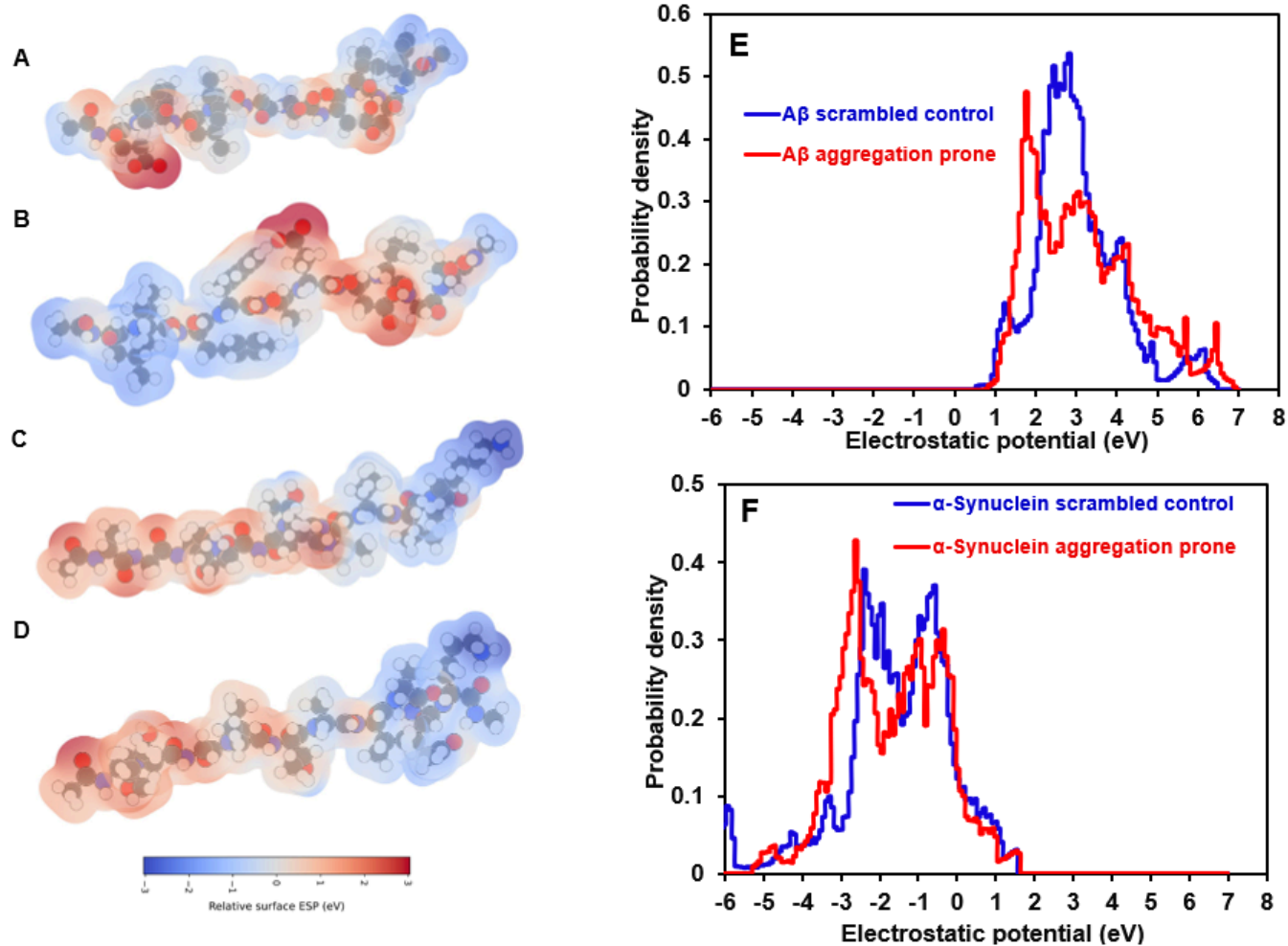


Figure 2

Electrostatic potential landscapes and surface ESP distributions of the peptide systems showing the ESP mapped molecular surfaces of (A) A β scrambled control, (B) A β aggregation prone peptide, (C) α -synuclein scrambled control, and (D) α -synuclein aggregation prone peptide. Relative surface ESP is shown from negative (blue) to positive (red) potential. Surface ESP probability density distributions compare the scrambled control and aggregation prone sequences for (E) A β and (F) α -synuclein.

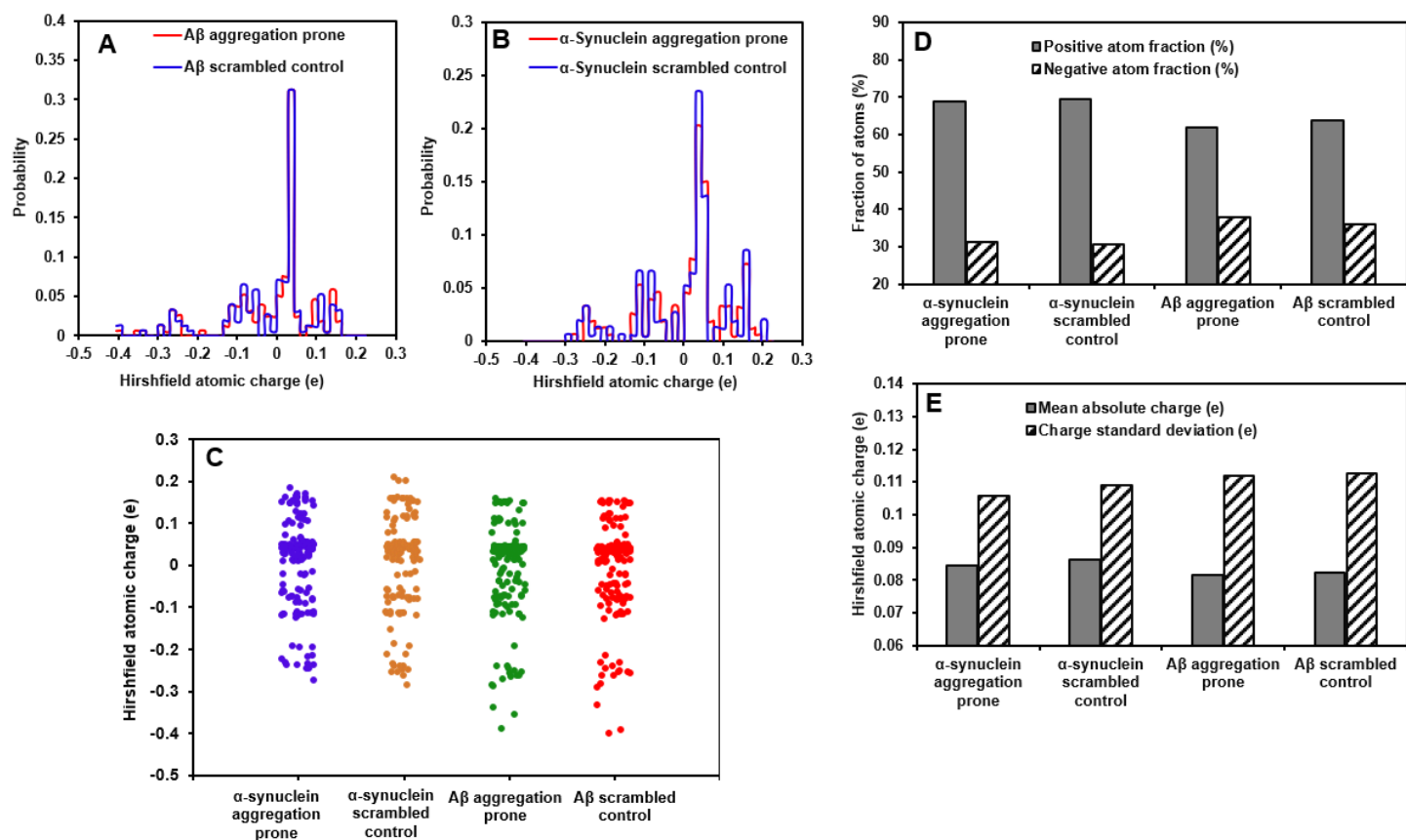


Figure 3

Hirshfeld atomic charge distributions of aggregation prone peptides and their composition matched scrambled controls for (A) atomic charge probability distributions for α -synuclein, (B) atomic charge probability distributions for $A\beta$, (C) individual atomic Hirshfeld charges for the four peptide systems and (D) fractions of atoms with positive and negative Hirshfeld charges.

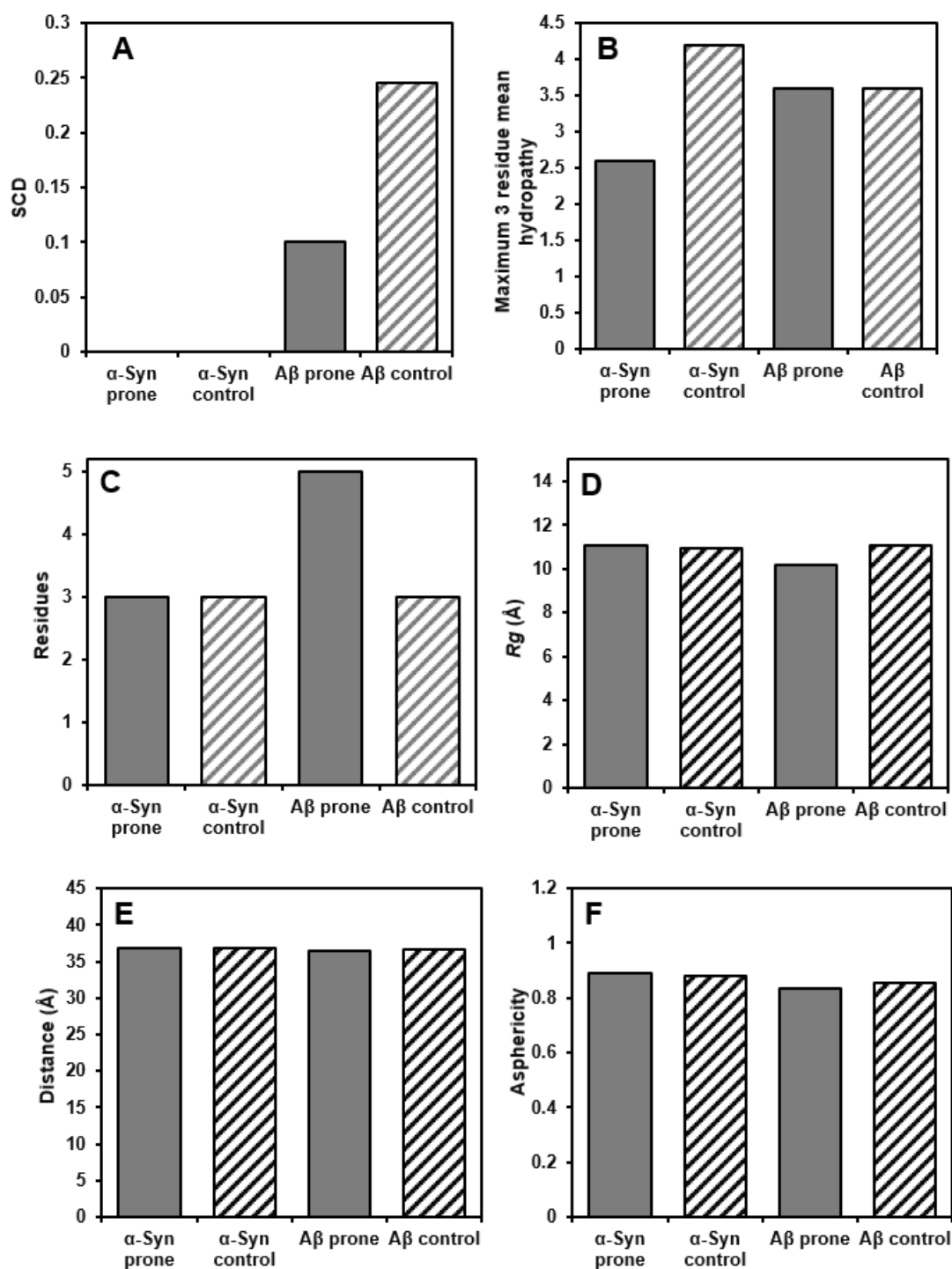


Figure 4

Sequence patterning, hydrophobic organization and molecular shape of aggregation prone and scrambled peptides showing (A) sequence charge decoration, (B) maximum three residue mean hydrophathy, (C) maximum hydrophobic run length, (D) R_g , (E) end-to-end distance and (F) molecular asphericity for aggregation prone α -synuclein and A β and their composition matched scrambled controls.

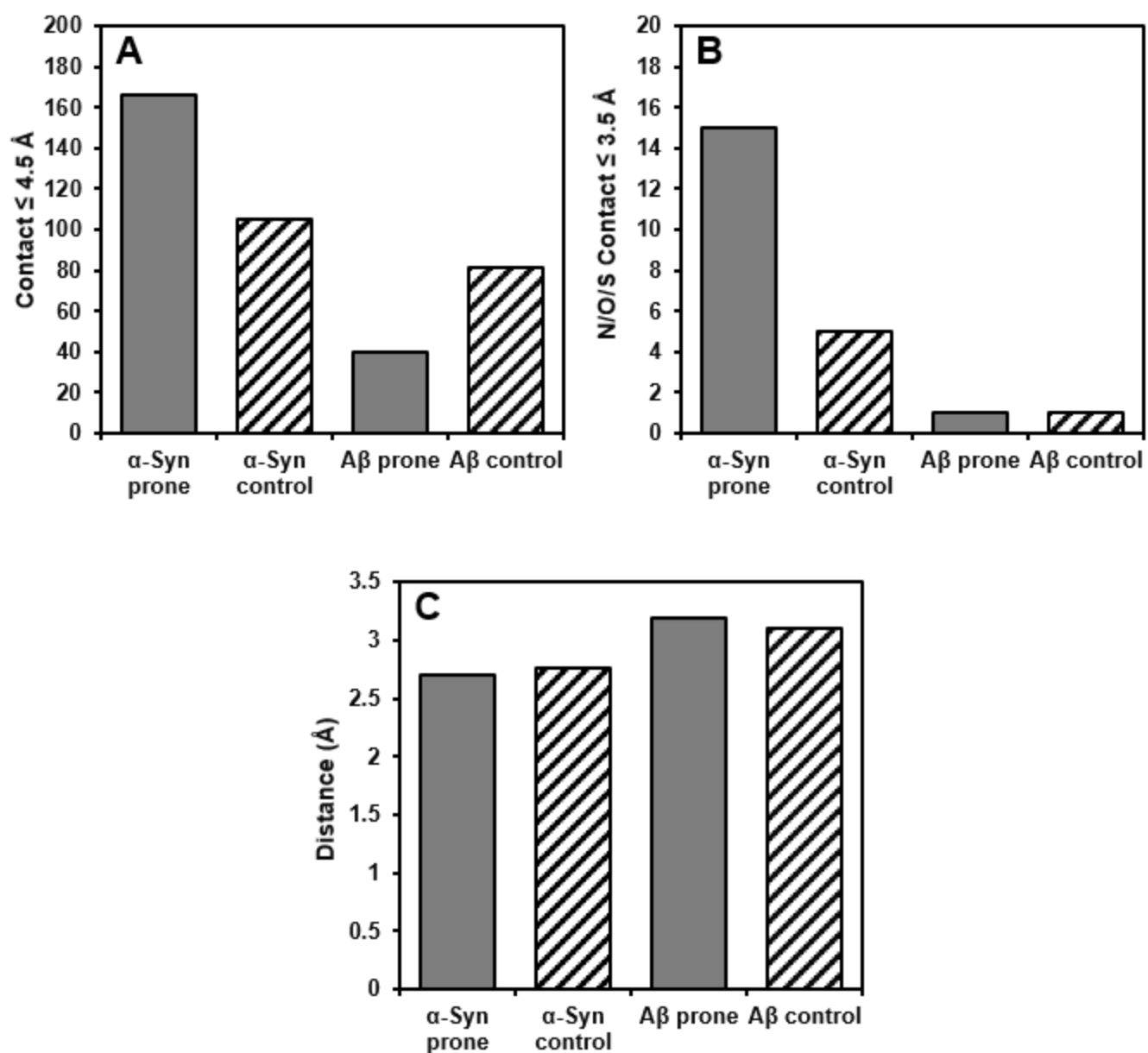


Figure 5

Intermolecular contact organization of aggregation prone and scrambled peptide dimers showing (A) number of intermolecular heavy atom contacts within 4.5 Å, (B) number of intermolecular N, O or S contacts within 3.5 Å and (C) minimum heavy atom intermonomer distance for aggregation prone α -synuclein and A β and their composition matched scrambled controls.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Scheme.1.png](#)