Protein folding, misfolding and aggregation: The importance of two-electron stabilizing interactions

Andrzej Stanisław Cieplak[1,2,3]*

1 Department of Chemistry, Bilkent University, Ankara, Turkey, **2** Department of Chemistry, Yale University, New Haven, Connecticut, United States of America, **3** Department of Chemistry, Brandeis University, Waltham, Massachusetts, United States of America

* cieplak@gmail.com

Abstract

Proteins associated with neurodegenerative diseases are highly pleiomorphic and may adopt an all-α-helical fold in one environment, assemble into all-β-sheet or collapse into a coil in another, and rapidly polymerize in yet another one via divergent aggregation pathways that yield broad diversity of aggregates' morphology. A thorough understanding of this behaviour may be necessary to develop a treatment for Alzheimer's and related disorders. Unfortunately, our present comprehension of folding and misfolding is limited for want of a physicochemical theory of protein secondary and tertiary structure. Here we demonstrate that electronic configuration and hyperconjugation of the peptide amide bonds ought to be taken into account to advance such a theory. To capture the effect of polarization of peptide linkages on conformational and H-bonding propensity of the polypeptide backbone, we introduce a function of shielding tensors of the C^α atoms. Carrying no information about side chain-side chain interactions, this function nonetheless identifies basic features of the secondary and tertiary structure, establishes sequence correlates of the metamorphic and pH-driven equilibria, relates binding affinities and folding rate constants to secondary structure preferences, and manifests common patterns of backbone density distribution in amyloidogenic regions of Alzheimer's amyloid β and tau, Parkinson's α-synuclein and prions. Based on those findings, a split-intein like mechanism of molecular recognition is proposed to underlie dimerization of Aβ, tau, αS and PrP^C, and divergent pathways for subsequent association of dimers are outlined; a related mechanism is proposed to underlie formation of PrP^{Sc} fibrils. The model does account for: (i) structural features of paranuclei, off-pathway oligomers, non-fibrillar aggregates and fibrils; (ii) effects of incubation conditions, point mutations, isoform lengths, small-molecule assembly modulators and chirality of solid-liquid interface on the rate and morphology of aggregation; (iii) fibril-surface catalysis of secondary nucleation; and (iv) self-propagation of infectious strains of mammalian prions.

Introduction

To understand protein folding, one needs to understand protein structure. And yet, in spite of the considerable interest and effort, even the most rudimentary issues of proteins conformational behaviour remain unresolved: 'Surprisingly, the field lacks a physicochemical theory of protein secondary structure' [1]. Indeed, for the chemist concerned to gain insight, protein study is in want of a theory which would explain what local backbone interactions make a residue a helix-breaker or a helix-former and how that propensity depends on the context—why a residue which is a helix-breaker in water becomes a helix-former in lipid or vacuum or why a residue which is a helix-former in the folding intermediate becomes a sheet-former in the native state. Pressing puzzles such as why certain sequences of helix-breakers and helix-formers may adopt an all-α fold in one environment, all-β fold in another, and collapse into a coil in yet another one, are bound to remain elusive unless these questions are answered.

The need to address the problem of secondary structure is underscored by the emerging evidence that 'folding is an inherently digital process in which the formative interactions are among backbone elements' [2], and that, given the folding and refolding of liquid lipase and gas-phase apomyoglobin, 'the exigency of water in determining protein folding could possibly be overstated in current models used to describe this phenomenon.' [3–6] To understand conformational behaviour of proteins, this evidence suggests, one ought to take into account the effects of main chain bonding—the backbone-backbone H-bonding but also the effects of hyperconjugative interactions of peptide linkages which depend on backbone conformation. The conceptual framework necessary for doing so is readily available in the modern physical organic chemistry [7]. The internal rotation in simple organic molecules was long argued to depend on the stereoelectronic as well as the steric and electrostatic effects [8,9]. The phenomena such as the gauche effect or the anomeric effect are now commonly attributed, at least in part, to hyperconjugation, and often satisfactorily described in terms of two-electron stabilizing interactions of the localized filled and vacant molecular orbitals [7–15]. Since the polypeptide chains comprise groups of orbitals whose through-bond and through-space interactions are altered by the internal rotation, similar effects may also play a role in the conformational equilibria of proteins. Hyperconjugative interactions of the peptide amide bonds have in fact been proposed by a number of authors to contribute to the stability of oligopeptides and proteins [16–22]. Thus, for the chemist to gain a sense of understanding, the theory of protein secondary structure may have to be rooted in the 'qualitative orbital thinking' [23,24] and the perturbational molecular orbital (PMO) theory which is uniquely equipped to support that way of thinking [7,25].

In this study, we attempt to construct such a theory. In a departure from the reigning paradigm, we take into account the variation in electronic configuration of the peptide bonds and the concomitant variation in the conformational and H-bonding propensity of the polypeptide backbone. The classical theory of protein structure was born half a century ago out of recognition of the steric and electrostatic constraints imposed by the peptide amide bonds on the internal rotation of the polypeptide chains [26,27]. The peptide linkages themselves, however, are in this theory assumed to be chemically equivalent and any differences in their geometry and bonding are disregarded. In contrast, we argue that the electronic configuration of the peptide amide bonds, and therefore the distribution of backbone density, does vary in the sequence and medium-dependent manner. This variation alters the balance between the backbone-backbone H-bonding and the polylactide-like backbone conjugation and thus underlies the diversity of conformational behaviour of the polypeptide chains.

The relationship between the electronic configuration of the peptide amide bonds and the torsional potential of the protein backbone was first considered in the context of modelling the

relay of chiral information via the π-conjugated systems [16]. Subsequently, a theory was proposed linking the secondary structure propensities of the Ala congeners (the Ala* amino acids) to the inductive and resonance effects of the side chains and the stereoelectronic effects [28–31]; the sense and relative importance of these effects was assumed to depend on the location of the peptide amide bonds along the amide rehybridization/polarization path [32]. To explore the theory's implications, a quantum-mechanical investigation of the single-site substitutions Ala→Lac (Lac≡L-lactic acid) was carried out [33].

We now return to the basic ideas espoused in these earlier studies, taking again advantage of the quantum-mechanical modelling of protein structure. However, we do not try to develop an *ab initio* or DFT simulation of folding—even if successful 'such an approach would enable one to mimic nature but not necessarily understand her' [34]. The PMO theory which underlies qualitative models of organic structure and reactivity, suggests the alternative approach: 'The PMO treatment is concerned with differences in the properties of structurally related molecules, rather than the absolute values of the quantities relating to the individual molecules. Actually, by an appropriate choice of the reference system it will prove possible to obtain the information we require directly, in a very simple and straightforward manner which preserves a close understanding of the chemistry involved' [35]. Following this protocol, we derive here a theory of encoding the 3D structure of proteins that attempts to address the origin of secondary structure and the mechanism of assembly of its elements into tertiary structure. This effort takes into account a wide range of phenomena such as divergent folding of highly homologous proteins and convergent folding of non-homologous proteins, metamorphic equilibria of lymphotactin, mitotic spindle protein Mad2 and *E. coli* virulence regulator RfaH, pH-driven transitions of viral fusion proteins and membrane translocation domains, acid-induced unfolding and fibrillization of transthyretin, synergistic folding of split inteins, and coupled folding and binding of molecular recognition features [36–39].

To be truly comprehensive, however, a theory of folding ought to address misfolding, and not just as a matter of the theory's elegance. One 'pressing puzzle' that presently tests our understanding of protein structure and folding is conformational behaviour of highly pleiomorphic proteins such as Aβ and tau proteins believed to play key role in genesis of Alzheimer's disease, α-synuclein (αS) associated with Parkinson's disease, or prion proteins (PrP) playing central role in genesis of transmissible spongiform encephalopathies (TSEs). Polymerization of Aβ, tau, αS or PrP presents especially demanding challenge given the divergence of aggregation pathways, broad diversity of aggregates' morphology, and apparent ease of proteins' transitioning through the entire secondary and supersecondary structure manifold. Therefore, to probe the explanatory and predictive power of our theory, we apply its tenets to develop a model for nucleate polymerization of these proteins. The model aims to account for the structural features of oligomers, non-fibrillar aggregates and fibrils—such as annular and tubular aggregation of paranuclei or the symmetry of protofilament assembly, and for the effects of incubation conditions, isoform length, point mutations, assembly modulators, surface catalysis, lipid-raft composition etc. In particular, we are interested in the effects of chirality on the rate and morphology of Aβ fibril formation on the surfaces of *R*(*S*)-cysteine-modified graphene oxide and self-assembled monolayers of *R*(*S*)-N-isobutyrylcysteine on gold [40,41], the differences between Aβ aggregation on hydrophilic mica and hydrophobic graphite [42], and the mechanism of Aβ fibril-surface catalysis of secondary nucleation [43]. Ultimately, however, our goal is to develop structure-based insights into the complexity of brain proteinopathies.

Computational methods

Our choice of the reference system for the PMO theory-informed analysis is a simplified structure of the polypeptide backbone which can be converted into a complete protein chain by 'switching on' three standard perturbations [44]: (1) the geometry perturbation (by changing conformation); (2) the atomic substitution/electronegativity perturbation (by plugging in the side chains); and (3) the intermolecular perturbation (by embedding backbone chain in polarizing or depolarizing environment and allowing for the interactions with co-solutes, other protein chains, surfaces of biopolyelectrolytes etc.). The simplified backbone structure comprises the $(-C^\alpha\text{-NH-C}(=O)-)_n$ chain and the localized MOs of the peptide amide bonds and their ligands. It is assumed that the sterically-allowed regions of the ψ/φ space are approximately equivalent in terms of energy unless the stereoelectronic and electrostatic interactions are introduced [27,45]. The relative importance of those interactions—which ones actually occur—depends on electronic configuration of the peptide amide bonds which in turn depends on electronic effects of the side chains (atomic substitution/electronegativity perturbation) and on the mutual polarization of the protein and the medium (intermolecular perturbation). Given that each interaction of the peptide amide bonds is maximized in a specific region of the ψ/φ space (geometry perturbation), the amino acid sequence and the medium may determine conformational and H-bonding propensity of the polypeptide backbone by 'switching on' and 'switching off' certain combinations of those interactions. To assess the effects of these perturbations, we employ (i) *ab initio* and DFT studies of secondary structure (geometry optimizations; NBO analysis of the donor-acceptor interactions of the localized natural bond orbitals using Weinhold's $\Delta E^{(2)}$ energies [12]—the BLW-ED energies [46] are lower but the qualitative trends are expected to be the same; SCRF modeling of solvent effects; GIAO calculations of the NMR shielding tensors $\boldsymbol{\sigma}(C^\alpha)^{Xaa}$ of C^α atoms), and (ii) qualitative concepts of two theories of solutions: the Onsager theory of solute-solvent polarization [47] and the Debye-Hückel theory of dilute solutions of strong electrolytes [48]. First, to assess the side chains' effect on the distribution of backbone density, the NMR shielding tensors $\boldsymbol{\sigma}(C^\alpha)^{Xaa}$ of the C^α atoms were calculated using the models of helix and sheet structures in gas phase (the effect of the medium is treated as a separate perturbation). Thus, two oligopeptides, N-acetyl hexaglycyl N-methylamide AcGGGGGGNHMe and N-acetyl pentaglycyl amide AcGGGGGNH$_2$, were initially optimized in the conformations corresponding to the hairpin with the type Ib reverse turn, and 3_{10}-helix, respectively, at the B3LYP/D95** level of the theory. The protocol involved folding of the peptide chain into the starting conformer using the standard φ_i and ψ_i values and subsequently an unconstrained optimization. The searches for the minima were completed by the default convergence criteria of Gaussian 98, Revisions A.3, A.7, A11.2 [49]. The sixth residue of the hexapeptide hairpin and the second residue of the pentapeptide helix were then systematically varied to generate congener structures with the canonical and covalently modified residues. The side chains were set into the conformations *trans* and–*gauche* about the C^α-C^β bond when needed, both in the neutral and ionized state when appropriate, and a number of structures were partially constrained as noted in S1 Table. Again, all the searches for the minima were completed by the default convergence criteria. The calculations yielded the total of 141 `AcGGGGGXaaNHMe` and `AcGXaaGGGNH`$_2$ structures. Atomic coordinates of the obtained structures were used to compute the NMR shielding tensors using the B3LYP/D95** and GIAO (Gauge-Independent Atomic Orbital) methods (S1 Table). The mean values of the obtained shielding tensors of the C^α atoms were converted into the folding constants σ^{Xaa} as described in Results and Discussion. To assess the effect of geometry perturbation on backbone conjugation and backbone-backbone H-bonding by examination of Weinhold energies $\Delta E^{(2)}$ of the hyperconjugative interactions which differentiate the

sterically-allowed regions of the Ramachandran map, the protocol described above was employed to optimize five models of secondary structure **1–5** (B3LYP/6-31G*): the decapeptide AcAAAAAAAAAANH$_2$ **1** in the 3$_{10}$-helix conformation, the pentadecapeptide AcAAAAAAAA AAAAAAANH$_2$ **2** in the α-helix conformation, the hexapeptide AcAAAAANHMe **3** in the 2$_7$-ribbon conformation, the ternary antiparallel complex of the tripeptide (AcAANHMe)$_3$ **4**, and the ternary parallel complex of the tripeptide (AcAANHMe)$_3$ **5**. In addition, the binary complexes of the oligopeptides (AcAAANHMe)$_2$ **6a**-**6d** and (AcAAAAANHMe)$_2$ **7a**-**7b**/**8a**-**8c** were obtained by unconstrained optimization (as described above, at the B3LYP/6-31G* level of the theory, completed by the default convergence criteria of Gaussian98). To assess the effect of a polar dielectric on the distribution of backbone density in a globular protein molecule, the model of TC5b: AcAAAAAAAAAGGPAAGAPPPA-NH$_2$ **9**, was obtained by the full unconstrained optimization (HF/3-21G, gas phase) of the peptide chain placed in the conformation defined by the φ/ψ angles reported for the NMR ensemble of TC5b [50]. The final gas-phase structure **9a** was re-optimized in water taken as the continuous dielectric (HF/3-21G, the Onsager model as implemented in the Gaussian suite with ε$_0$ = 78.39 and the radius of the spherical solvent cavity a$_0$ = 8.43 Å) until the default convergence criteria were fully met again to obtain the structure **9b**. The Cartesian atomic coordinates and the total energies of all the secondary and tertiary structure models (156 entries) are included in S1 Dataset.

Results and discussion

a. The PMO theory of the secondary and tertiary structure of proteins

(i) Electronic configuration of the peptide amide bonds and conformational and H-bonding propensity of the polypeptide backbone.   The question how the two-electron stabilizing interactions relate to stability of protein structure can be addressed by considering how backbone hyperconjugation and covalent contributions to backbone-backbone H-bonding vary in the sterically allowed regions of the ψ/φ space, cf. Fig 1A. To maximize secondary orbital overlap [10], backbone hyperconjugation involves primarily pairs of adjacent peptide amide bonds where the 'upstream' *i−1/i* bond is a donor and the 'downstream' *i/i+1* bond is an acceptor; the change in their mutual orientation, the 'geometry perturbation', changes the nature and magnitude of the interaction. When the planes of these two bonds are approximately perpendicular, φ$_i$ = −90˚±30˚, the fold ensures optimal orbital overlap for the generalized anomeric effect [51,52] and the covalent contributions to the N$_i$···C$_i$ = O and C$_{i-1}$ = O···C$_i$ = O interactions, see Fig 1B [17–20,33,53–58]; in contrast, orbital overlap for the covalent contributions to backbone-backbone H-bonding is relatively poor [59]. When these two peptide bonds are approximately coplanar, φ$_i$ = −150˚±30˚, the situation is reversed. The generalized anomeric effect and homohyperconjugation are diminished due to vanishing overlap and the significance of the alternative hyperconjugative stabilization is limited [21] while the overlap for covalent contributions to backbone-backbone H-bonding is optimal [59]. The estimated energies Δ$E^{(2)}$ of the corresponding donor-acceptor interactions in the secondary structure models **1**-**5**, cf. Computational Methods, are listed in S2 Appendix.

It follows that the peptide bonds which are good N lp donors and poor H-bond acceptors should stabilize the φ$_i$ = −90˚±30˚ fold while the peptide bonds which are poor N lp donors and good H-bond acceptors should stabilize the φ$_i$ = −150˚±30˚ fold. To refine this picture we take into account computational and experimental evidence of wide variation in electronic configuration of the amide bonds in carboxamides, lactams, oligopeptides and proteins [7,32,60–63]. According to this evidence, one can describe bonding of peptide linkages in terms of varying contributions of the five resonance structures I-V, see Fig 1C. The shift I→II→III→IV→V has the following consequences for the ψ/φ space preferences:
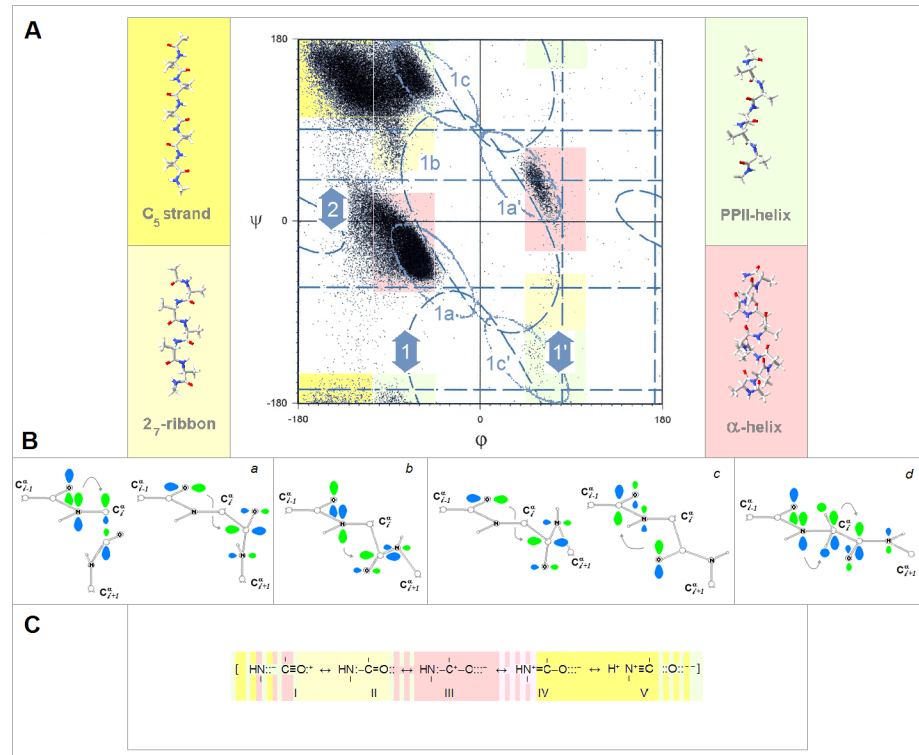
Fig 1. Electronic configuration of the peptide amide bonds and conformational and H-bonding propensity of the polypeptide backbone. (A) The dependence of interactions of the adjacent peptide amide bonds on location in the ψ/φ space of the polypeptide backbone (the diagram is adapted from the ref. [45]). The variation in φ_i changes mutual orientation of the peptide bonds: the bond planes are approximately perpendicular to each other in the helical region **1** ($\varphi_i = -90°\pm30°$) and approximately coplanar in the extended strand region **2** ($\varphi_i = -150°\pm30°$). The variation in ψ_i changes the extent of backbone H-bonding in the helical sub regions **1a–1c**. **(B)** Two-electron stabilizing interactions of the peptide amide bonds, depicted using the canonical amide MO's: (a) the generalized anomeric effect $\pi_2(N_i–C'_{i-1} = O)\rightarrow\sigma*(C^\alpha–C'_i)$ which is maximized when the $C^\alpha_i–C'_i$ bond, the best hyperconjugative σ acceptor at C^α_i, overlaps the N_i lp that is in the entire helical region **1** ($\varphi_i = -90°\pm30°$), and homohyperconjugation $n(C'_{i-1} = O)\rightarrow\pi_3*(N_{i+1}–C'_i = O)$ maximized in the α-helix region **1a** ($\psi_i = -30°\pm30°$); (b) homohyperconjugation $\pi_2(N_i–C'_{i-1} = O)\rightarrow\pi_3*(N_{i+1}–C'_i = O)$, maximized in the 2_7-ribbon (C_{7eq} and C_{7ax}) region **1b** ($\psi_i = 90°\pm30°$); (c) homohyperconjugation $n(C'_{i-1} = O)\rightarrow\pi_3*(N_{i+1}–C'_i = O)$ and $n(C'_i = O)\rightarrow\pi_3*(N_i–C'_{i-1} = O)$, maximized in the PP$_{II}$-helix region **1c** ($\psi_i = 150°\pm30°$); (d) the extended (double) hyperconjugation $\pi_2(N_i–C'_{i-1} = O)\rightarrow\pi(C^\alpha_iRR')\rightarrow\pi_3*(N_{i+1}–C'_i = O)$ maximized in the C_5 region **2** ($\varphi_i = -150°\pm30°$). **(C)** Modern resonance model of the amide bonding and the dependence of conformational and H-bonding propensity of the polypeptide backbone on electronic configuration of the peptide amide bonds, see the text section **a**.(i).

https://doi.org/10.1371/journal.pone.0180905.g001

(1) The structure I contributes to the configuration of the least-polarized bonds which display positive $r_{C = O}$ vs. r_{C-N} correlations [32], are relatively poor acceptors of H-bonds, form largely ionic backbone-backbone H-bonds [64,65], and are good π/resonance N lp donors; it is compatible with the PP$_{II}$ helix where backbone-backbone H-bonding is absent as well as the $\varphi_i = 180°\pm30°/\psi_i = 180°\pm30°$ fold cf. Fig 1B(d) which is stabilized by the extended N lp hyperconjugation [21]. The least polarized backbone segment is therefore expected to exist as a statistical random coil unless molecular embedding turns it into an extended strand (the C_5^* strand) or a helix (PP$_{II}$-helix or the α^*-helix).

(2) The structure II contributes to the configuration of the moderately-polarized bonds which are still relatively poor acceptors of H-bonds but good N lp donors. This configuration is compatible with the 2_7-ribbon (the C_{7eq} strand) since the $\varphi_i = -90°\pm30°/\psi_i = 90°\pm30°$ fold is

stabilized by the generalized anomeric effect as well as the homohyperconjugation, cf. Fig 1B (a) and 1B(b), and by the relatively effective backbone-backbone H-bonding. Thus, the moderately-polarized segments (largely II) can form β-sheets via assembly of the C_{7eq} strands as reported in Fig 2.

(3) The structure III contributes to the configuration of the polarized bonds which are still good N lp donors and already good H-bond acceptors while being good C' σ and π acceptors as well. Thus, this configuration is unique in ensuring that the three major interactions which stabilize the α-helix are sufficiently strong at the same time (the general anomeric effect and
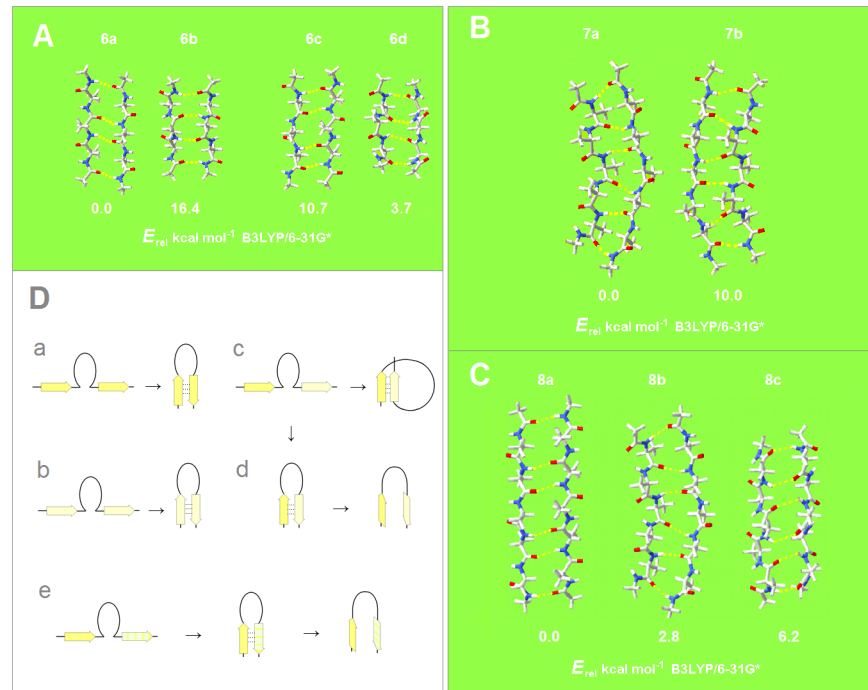
Fig 2. Conformational diversity in the binary complexes of extended oligopeptide strands.
The geometries and energies obtained by quantum-mechanical modeling of the two-stranded β-sheets **6–8** (Computational Methods). The individual strands in these complexes optimize either to the C_5 or the C_{7eq} (2_7-ribbon) geometries, and their conformations are same in the antiparallel complexes ($C_5\uparrow C_5\downarrow$ or $C_{7eq}\uparrow C_{7eq}\downarrow$) and mixed in the parallel complexes ($C_{7eq}\uparrow C_5\uparrow$); *the antiparallel complexes with mixed strand conformations ($C_{7eq}\uparrow C_5\downarrow$) are unstable in unconstrained optimizations*. **(A)** The antiparallel complexes of the tetrapeptides (AcNH-Ala$_3$-NH$_2$)$_2$ displaying the edge-to-edge topoisomerism: the assembly creates either two or one large H-bonded (HB) ring. **1a**: the $C_5\uparrow C_5\downarrow$ complex with two large HB rings; **1b**: the $C_{7eq}\uparrow C_{7eq}\downarrow$ complex with one large HB ring; **1c**: the $C_5\downarrow C_5\uparrow$ complex with one large HB ring; **1d**: the $C_{7eq}\downarrow C_{7eq}\uparrow$ complex with two large HB rings. **(B)** The parallel complexes of the hexapeptides (AcNH-Ala$_5$-NH$_2$)$_2$ displaying the edge-to-edge topoisomerism: here all the H-bonded rings are equivalent but complex formation involves the edges with either two or three intrachain H-bonds. **2a**: the $C_{7eq}\uparrow C_5\downarrow$ complex involving the edges with two intrachain H-bonds; **2b**: the $C_5\downarrow C_{7eq}\downarrow$ complex involving the edges with three intrachain H-bonds. The large difference in the energy of the edge-to-edge topoisomers is not observed in the case of the binary complexes of the oligopeptides with the odd number of the peptide bonds. **(C)** The relative energies of the **3a**: $C_5\downarrow C_5\uparrow$, **3b**(≡**2a**): $C_{7eq}\uparrow C_5\downarrow$, and **3c**: $C_{7eq}\uparrow C_{7eq}\downarrow$ complexes of the hexapeptides (AcNH-Ala$_5$-NH$_2$)$_2$. **(D)** The segments comprising two consecutive strands form stable β-hairpins (antiparallel assembly) when the two strands are either (a) both highly polarized ($C_5\uparrow C_5\downarrow$) or (b) both moderately polarized ($C_{7eq}\uparrow C_{7eq}\downarrow$) (color-coding as in Fig 1). In contrast, when one strand is highly polarized and the other is moderately polarized, these segments are expected to form (c) β-solenoid coils (parallel assembly, $C_{7eq}\uparrow C_5\uparrow$) or (d) unstable β-hairpins (antiparallel assembly $C_{7eq}\uparrow C_5\downarrow$) which are prone to convert into β-arches; similarly (e) when one strand is highly polarized and the other is least-polarized (the configuration described by a large contribution of the structure I, Fig 1C), the segment may form a hairpin ($C_5\uparrow C_5*\downarrow$) which is also prone to convert into β-arch.

https://doi.org/10.1371/journal.pone.0180905.g002

homohyperconjugation, Fig 1B(a), and the backbone-backbone C = O· · ·H−N bonding). The polarized segments (largely III) are therefore expected to form α-helices.

(4) The structures IV and V contribute to the configuration of the highly- and most-polarized bonds which display negative $r_{C = O}$ vs. r_{C-N} correlations [32], are good acceptors and donors of H-bonds and form largely covalent H-bonds (including the backbone-backbone C = O· · ·H−C bonding [66,67]), but are poor π donors in the hyperconjugative interactions of N. Thus, the highly-polarized backbone segments will stabilize the $\varphi_i = −150°±30°/\psi_i = 150°±30°$ fold i.e. the C_5 strands which readily assemble into β-sheets as reported in Fig 2. However, the most-polarized segments may also stabilize the PP_{II}-helix and turn folds, *vide infra*, via the homohyperconjugation cf. Fig 1B(a) and 1B(c), and thereby destabilize the C_5 strands.

This analysis implies that *the contribution of the energy of two-electron stabilizing interactions ($\Delta E^{(2)}$stabilization [12]) to $\Delta G_{coil\rightarrow helix}$ has one minimum with respect to charge polarization of the polypeptide backbone while the contribution to $\Delta G_{coil\rightarrow sheet}$ has two such minima.* Assuming that the $\Delta E^{(2)}$ contributions are significant, one expects $\Delta G_{coil\rightarrow helix}$ and $\Delta G_{coil\rightarrow sheet}$ to be quadratic and quartic functions, respectively, of backbone polarization.

According to the data for the models of β structure **6**-**8** in Fig 2, the preferred mode of assembly of the two-stranded β-sheets may also depend on charge polarization of the main chain. The antiperiplanar assembly should be stabilized when two strands are either both highly polarized, $C_5\uparrow C_5\downarrow$, or both moderately polarized, $C_{7eq}\uparrow C_{7eq}\downarrow$ (backbone-polarization 'symmetry'); the parallel assembly should be stabilized when one strand is highly polarized and the other is moderately polarized, $C_5\downarrow C_{7eq}\downarrow$ (backbone-polarization 'asymmetry'). This model implies that the segments comprising two consecutive strands form stable β-hairpins (antiparallel assembly) when the two strands are either both highly polarized ($C_5\uparrow C_5\downarrow$), Fig 2D(a), or both moderately polarized ($C_{7eq}\uparrow C_{7eq}\downarrow$), Fig 2D(b). In contrast, when one strand is highly polarized and the other is moderately polarized, these segments are expected to form β-solenoid coils (parallel assembly, $C_5\uparrow C_{7eq}\uparrow$), Fig 2D(c), or unstable β-hairpins (antiparallel assembly $C_5\uparrow C_{7eq}\downarrow$) which may convert into β-arches, Fig 2D(d). Lastly, when one strand is highly polarized and the other is least-polarized, the segment may form a hairpin ($C_5\uparrow C_5^*\downarrow$, see Fig 1C) also prone to convert into β-arch, Fig 2D(e). Thus, in addition to β turns, β bulges, β arcs and small H-bonded rings [68], it is the number, spacing and sequence of the C_5, C_5^* and C_{7eq} segments that may direct the polypeptide chains to fold into β structure of specific chirality and topology such as β meanders, up-and-down β barrels, Greek-key motifs and β-roll barrels, β sandwiches, β solenoids or β arcades.

This analysis also suggests that the contribution of the energy of two-electron stabilizing interactions to $\Delta G_{coil\rightarrow turn}$ correlates with the *change* in charge polarization along the polypeptide backbone. The juxtaposition of the polarizing and depolarizing residues maximizes the $C_{i-1} = O· · ·C_i = O$ and $C_i = O· · ·C_{i-1} = O$ interactions (the homohyperconjugation), Fig 1B(a) and 1B(c). Consequently, a steep decrease in charge polarization along a short segment of the polypeptide backbone is expected to stabilize the elements of secondary structure where such interactions are likely to play a role: the 3_{10}- and PP_{II}-helices as well as β turns, $\alpha_R\alpha_L$ strands, classic β bulges, and β spirals i.e. collagen and elastin among others. The large change in backbone polarization can be achieved either by introduction of the least-polarized segment (with a large contribution of the structure I) or by introduction of the most polarized segment (with a large contribution of the structure V). Thus, β turns are expected to have two distinct electronic markers similarly to β strands.

(ii) Primary sequence and the intrinsic pattern of polarization of the peptide amide bonds: Folding potential *FP* of the polypeptide backbone. The intrinsic pattern of charge polarization of the polypeptide backbone of a given protein is determined by the primary

sequence as a result of the steric, field/inductive, and resonance effects of the side chains operating in the immediate vicinity of the C^α atoms—the 'atomic substitution/ electronegativity perturbation'. The NMR shielding tensors $\sigma(C^\alpha)^{Xaa}$ of the C^α atoms, calculated using the hairpin with the type Ib reverse turn and the 3_{10}-helix as the models (the L-amino acid series, 141 entries in S1 Table) at the B3LYP/D95** level of the theory (see the Computational Methods), are taken as a measure of the cumulative effect of these interactions on the distribution of backbone density. The folding constants σ^{Xaa}, listed in Table 1, are derived from the linear normalization of the $\sigma(C^\alpha)^{Xaa}$ tensor values to the scale where the σ^{Pro} constant for proline is −1 and the σ^{Gly} constant for glycine is 1. The amino acid residues can thus be said to be polarizing when $\sigma^{Xaa} < 0$ and depolarizing when $\sigma^{Xaa} > 0$; note that the polarizing effect depends on the ionization state of the side chains i.e. on the polarity and pH of the medium.

Using the folding constants σ^{Xaa}, we quantify the relationship between the side chains' effect and the conformational and H-bonding propensity of the polypeptide backbone by the magnitude and the slope of *the folding potential FP*. The folding potential at the residue i, FP_i, is defined as the averaged sum of the mean μ_i and standard deviation σ_i of the constants σ^{Xaa} within the three-(i−1, i, i+1) and five-(i−2, i−1, i, i+1, i+2)-residue windows, Eq. (1). The folding constants σ^{Xaa} are averaged over two windows of different width to account for the neighbouring residue effect, and the standard deviation terms are added to account for the effect of juxtaposing the polarizing and depolarizing residues (the weight of the μ_i and σ_i terms is arbitrary at this point): $FP_i = \frac{1}{2}[\mu_i(\sigma^{Xaa}_j; j = i−1, i, i+1) + \sigma_i(\sigma^{Xaa}_j; j = i−1, i, i+1) + \mu_i(\sigma^{Xaa}_j; j = i−2, i−1, i, i+1, i+2) + \sigma_i(\sigma^{Xaa}_j; j = i−2, i−1, i, i+1, i+2)]$ (1).

The slope of the folding potential at the residue i, $\Delta FP_{i-1 \rightarrow i+1}$, is approximated by the difference of the folding potential at the residues i−1 and i+1: $\Delta FP_{i-1 \rightarrow i+1} = FP_{i+1} − FP_{i-1}$ (2). This definitions imply that proteins may tolerate the 'inverted' sequences [69].

**Table 1. Folding constants σ^{Xaa} of the canonical amino acids:[a,b].**

Xaa	σ^{Xaa}	Xaa	σ^{Xaa}
A	0.1898	K	−0.0772
C	−0.4989	L	−0.0441
C[SMe]	−0.0403	M	−0.2143
D	0.1293	N	0.0296
D⁻	−0.1087	P	−1
E	0.1889	Q	−0.2485
E⁻	−0.4847	R	0.1683
F	−0.4289	S	−0.4700
G	1	T	−0.9066
H	−0.2917	V	−0.7703
H⁺	0.2584	W	−0.2704
I	−0.7647	Y	−0.3981

[a] Each tensor $\sigma(C^\alpha)$ is the average of the values obtained with two models of secondary structure, a hairpin (AcGGGGXGNHMe/Ib) and a helix (AcGXGGGNH$_2$/3_{10}), at the GIAO//B3LYP/D95** level of the theory. The mean values for the *trans* and −*gauche* conformers of the side chain about the C^α-C^β bond are taken when appropriate.

[b] The σ^{Xaa} constants for some covalently modified amino acids are: Ser O^γ-PO$_3^{-2}$ −1.1584, Met S (=O)$_2$ −0.1101, Lys N$^\xi$-COCH$_3$ −0.2518, Val C^β-CH$_3$ −0.9993, Ala C^β-F$_3$ −0.4258, Phe -F$_5$ −0.0829, Leu C^δ-F$_3$/C^δ-F$_3$ 0.0049, Ala C^β-CF$_3$ 0.2713, Ala C^β-n-CH$_2$CH$_2$CH$_3$ −0.2014, Thr C^β-N$^\gamma$H$_2$ −0.8477, Gly C^α-C≡N 0.8893, Gly C^α-C≡NO 0.8500, Gly C^α-C≡CH 0.6828.

$\sigma^{Xaa} = \{[\sigma(C^\alpha)^{Xaa}(trans) + \sigma(C^\alpha)^{Xaa}(−gauche)] − [\sigma(C^\alpha)^{Gly} + \sigma(C^\alpha)^{Pro}]\} / [\sigma(C^\alpha)^{Gly} − \sigma(C^\alpha)^{Pro}]$.
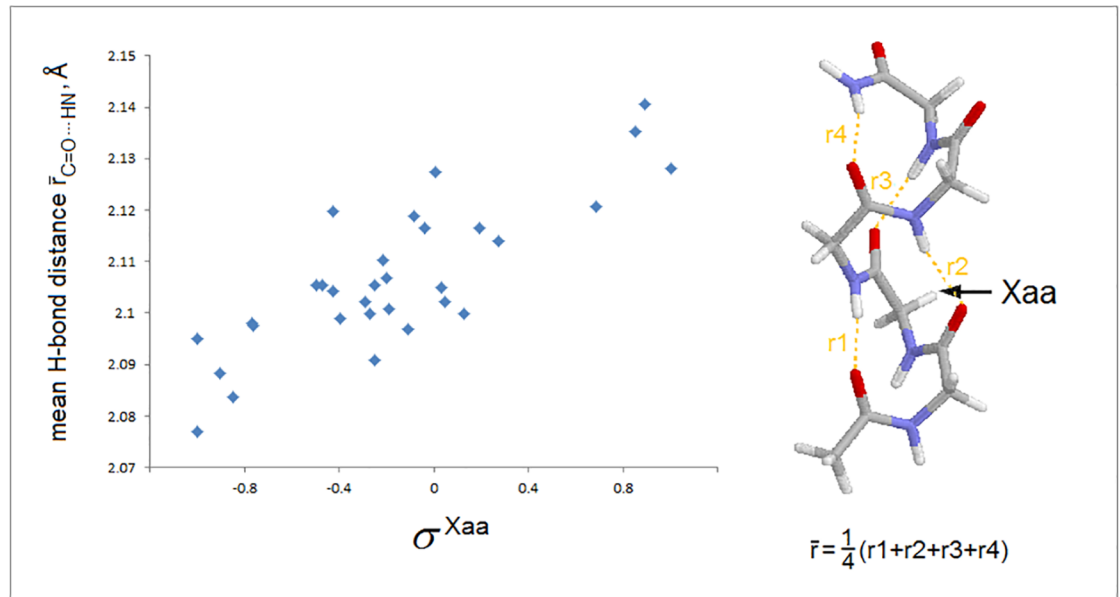
Fig 3. Folding constants σ^{Xaa} and the energy of backbone-backbone H-bonding. The calculated average backbone-backbone H-bond distance in the 3$_{10}$-helices AcNHG(Xaa)GGGNH$_2$ (shown in the right hand panel, calculated at the B3LYP/D95** level of the theory, cf. Computational Methods) vs. the folding constants σ^{Xaa} for all except the ionized Xaa residues listed in Table 1.

https://doi.org/10.1371/journal.pone.0180905.g003

As shown in Fig 3, the folding constants σ^{Xaa} account for a significant fraction of variation in the averaged distances of backbone-backbone C = O···H−N bonds in the AcNHG(Xaa)GGGNH$_2$ 3$_{10}$-helices (calculated at the B3LYP/D95** level of the theory, cf. Computational Methods). The correlation confirms that the σ^{Xaa} constants, and hence the folding potential FP_i, provide a measure of the intrinsic pattern of backbone polarization of a given protein.

The examples of the FP_i plots for the small soluble domains in Fig 4 suggest that the sequences which support spontaneous formation of the archetypal elements of secondary structure are marked by the specific values of the folding potential and the specific patterns of its slope. For instance, for the solvent-exposed α-helices, the $\Delta G_{coil \to helix}$ minimum seems to occur in the range $0 < FP_i < 0.3$.

The expected $\Delta FP_{i-1 \to i+1}$ patterns for the archetypal 'helix', 'strand' and 'turn', and the corresponding FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots are shown in Fig 5, along with the illustrative examples of such plots obtained for the autonomously folding models of β structure [71–75]. As discussed earlier, the archetypal 'strand' and 'turn' elements have each two avatars: (i) the 'C$_5$ strand' (highly-polarized segment, cf. structure IV in Fig 1) and the 'C$_{7eq}$ strand' (moderately-polarized segment, cf. structure II in Fig 1, see also Fig 2), and (ii) the '$FP_i \gg 0$ turn' (least-polarized segment, cf. structure I in Fig 1) and the '$FP_i \ll 0$ turn' (most-polarized segment, cf. structure V in Fig 1). Since the optimal FP_i values for each secondary structure element depend on the medium's capacity to polarize the protein, *vide infra*, the ordinates of the characteristic clusters in Fig 5 will change with the environment, including microenvironment of molecular embedding.

(iii) Polarization of the polypeptide backbone by the medium: Folding basin's gradient of permittivity and organization of tertiary structure. In addition to the side chains' impact, electronic configuration of the polypeptide backbone is affected by the mutual polarization of the protein and its environment—the 'intermolecular perturbation' [47]. The mutual
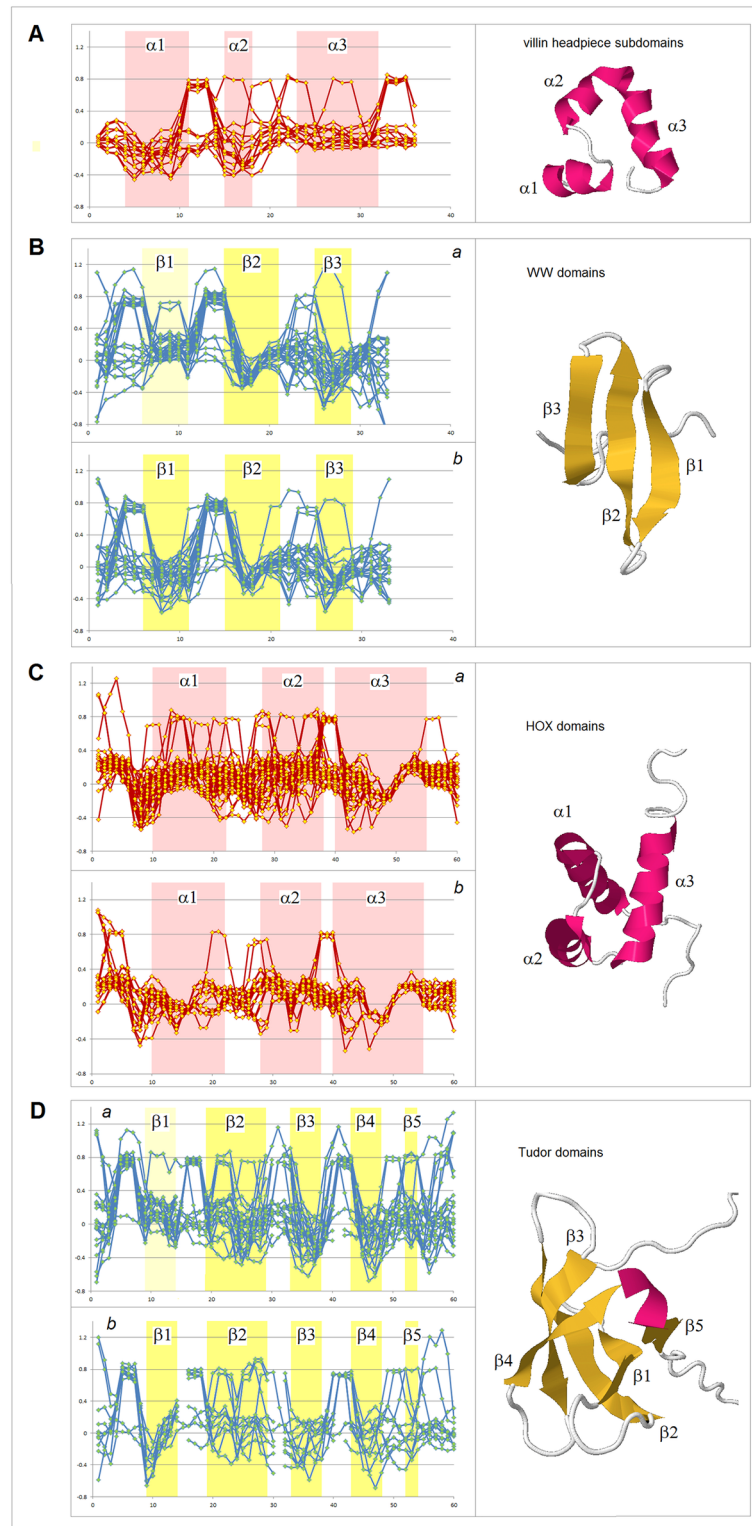
Fig 4. FP_i plots for small all-α and all-β soluble proteins. The folding potential at the residue i (FP_i, calculated from Eq. (1)), is plotted (Y-axis) against the residue number i (X-axis). The multiple alignments are taken from the SMART database (smart.embl-heidelberg.de) and the reference below. Note the characteristic FP_i profiles of the secondary structure elements and the variation in average FP_i values of those elements, $\overline{FP_i}$(α) or $\overline{FP_i}$ (β): **(A)** VHP (villin headpiece) domain, accession SM00153; **(B)** WW domain, accession #

SM00456: (a) $\overline{FP_i}$ (β1) > 0; (b) $\overline{FP_i}$ (β1) < 0; **(C)** HOX (homeobox) domain: (a) $\overline{FP_i}$ (α1) > 0; (b) $\overline{FP_i}$ (α1) < 0 [70]; **(D)** Tudor domain, accession SM00333: (a) $\overline{FP_i}$ (β1) > 0; (b) $\overline{FP_i}$ (β1) < 0. The helical and extended segments of the protein chain are shaded red, and yellow or light yellow, respectively.

polarization with a continuous dielectric engages peptide bond dipoles as well as molecular electric moments e.g. the helix macrodipole. The interaction results, among others, in the change in the free energy barrier to internal rotation about the amide C-N bond; this change was shown to correlate with the dielectric constant function $(\varepsilon-1)/(2\varepsilon+1)$ [76]. This correlation implies that the helix or cross-β arrays of H-bonded peptide linkages become more polarized upon the transfer from a non-polar to a polar medium, and our *ab initio* study of the model of TC5b mini-protein [50] supports this conclusion, see Fig 6 and the Computational Methods.

Thus, the polarization of those H-bonded networks is expected to increase with the increase in relative permittivity of the surrounding medium in the order: gas phase, lipid matrix of the phospholipid bilayer, the interior of a globular protein or DNA duplex [77,78], the interface of a phospholipid bilayer, a micellar interface, nematic phase of unspun silk, 'Teflon-coating' of the polypeptide chain in dilute water/(TFE or HFIP) solutions [79], the ~4 M in KCl cytosol of extreme halophilic *Archea* [80], and the cytosol or blood serum under the standard physiological conditions. The screening effect of the medium on the coulombic contribution to backbone-backbone H-bonding may be important in the case of the least-polarized backbone segments that are akin to the low-MW secondary amides in terms of electronic structure. Both experimental and computational evidence suggest that the enthalpy of H-bonding between such amides goes nearly to zero in water [64,65]. However, covalent contribution to H-bonding cannot be neglected even in the case of the water dimer in liquid water [81]. It seems reasonable to expect that the screening effect is negligible in the case of largely covalent backbone-backbone H-bonding of the polarized peptide bonds [12].

It follows that the secondary structure propensity is a function of both the intrinsic pattern of main-chain polarization, defined here by the folding potential FP_i, and the capacity of the environment to polarize the polypeptide backbone. The expected trend is shown in Fig 7: as the polarizing capacity of the environment increases on going from vacuum and lipids to aqueous buffers and cross-β structure, the values of FP_i which are optimal for the stability of a given element of secondary structure become more positive. For instance, the free energy $\Delta G_{\text{coil}\rightarrow\text{helix}}$ has one minimum with respect to the folding potential, cf. the FP_i region color-coded red in the diagram. It is expected that a negative value of FP_i is required to ensure helix stability in nonpolar environments, Fig 7(a) and 7(b), and that upon the transfer to an aqueous medium the position of the $\Delta G_{\text{coil}\rightarrow\text{helix}}$ minimum shifts to a positive value of FP_i, Fig 7(c).

We propose that the tendency to maintain congruity of the folding potential, medium and secondary structure drives the organization of the tertiary structure. Upon the collapse of the polypeptide chain or its segment into a compact conformation, the emerging shell→core depression or elevation of dielectric permittivity generates *the folding basin FB*. The position of an element of secondary structure within the folding basin, i.e. either in the interior or on the surface of the compact structure, depends on the deviation of its folding potential from the optimal 'helix' or 'strand' value. For instance, in the aqueous buffers, the helix with the more *negative* than optimal folding potential, e.g. incorporating the 'C₅ strand' segment, will be stabilized when it is buried in the compact structure's interior which has lower relative permittivity than the bulk of the solvent (the folding basin is a depression of relative permittivity with respect to the aqueous buffer). On the other hand, the helix with the more *positive* than optimal folding potential, e.g. incorporating the ' $FP_i \gg 0$ turn' or 'C₇ₑq strand' segment, will anchor
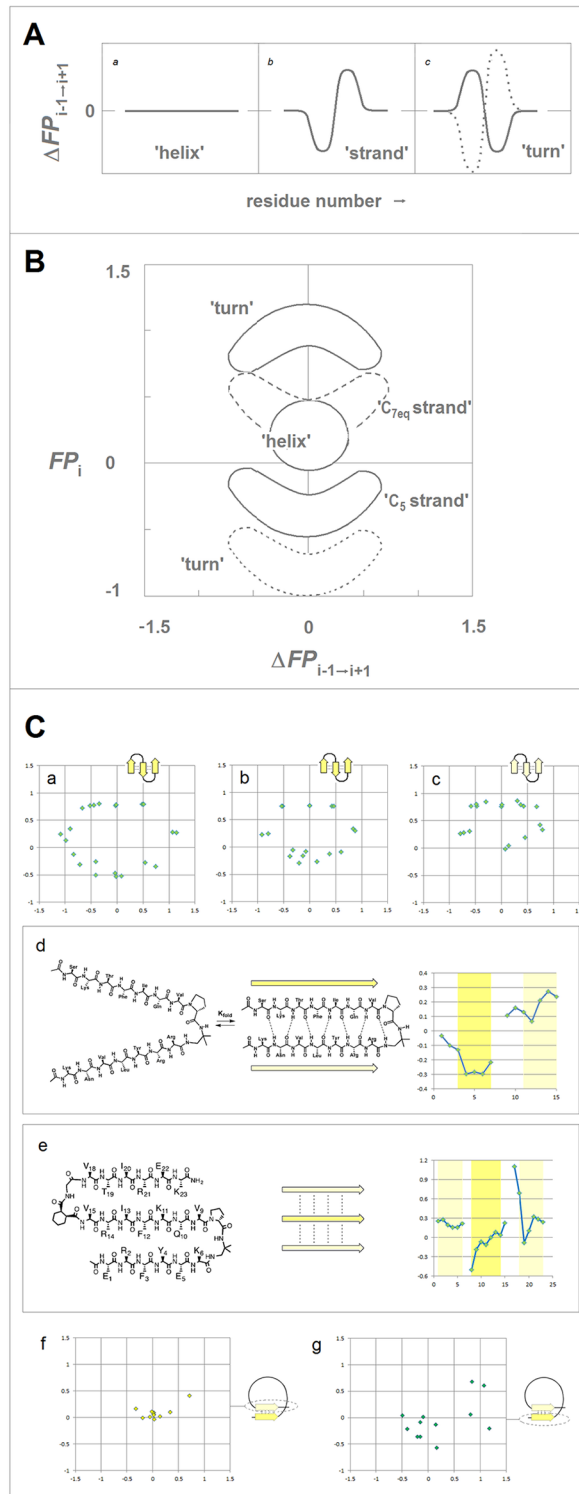
Fig 5. FP_i as a probe of the three-dimensional structure of proteins. (A) The patterns in the plots of $\Delta FP_{i-1\rightarrow i+1}$ (Eq. 2) vs. the residue number, characteristic of the archetypal 'helix', 'strand' and 'turn'. **(B)** Characteristic clusters of the data sets in the plots of FP_i vs. the 'slope' of FP_i, $\Delta FP_{i-1\rightarrow i+1}$, which correspond to the three archetypal elements of the secondary structure: e.g. the presence of the archetypal 'helix' will be marked by a compact cluster of data sets in the center of the plot. The ordinate of this cluster will vary since the optimal FP_i value for 'helix' depends on the medium's capacity to polarize the protein, *vide infra*. Note that 'strand' and 'turn' have each two avatars: (i) 'C_5 strand' and 'C_{7eq} strand', and (ii) ' FP_i>>0 turn' (defined here

as the three- or five-residue segment that incorporates Gly in the centre) and 'FP_i<<0 turn'. **(C)** The presence of the archetypal antiparallel 'sheet' would be marked by a circular distribution of data sets that combines the 'C$_5$ strand'/'turn' or 'C$_{7eq}$ strand'/'turn' clusters while the presence of the parallel 'sheet' would be marked by a combination of the 'C$_5$ strand' and 'C$_{7eq}$ strand' clusters, cf. Fig 2. This is illustrated by examples of *de novo* designed three-stranded antiparallel β-sheets (three-stranded β meanders), two- and three-stranded parallel β-sheets, and two-stranded parallel β-sheets embedded in left-handed coils from the C-terminal domains of the penicillin binding protein PBP2x from *Streptococcus pneumoniae*, PDB ID 1k25: (a) KGEWTFVNGKYTV SINGKKITVSI, ~50% in β structure, H$_2$O, pH 3, 25˚C (C$_5$↑C$_5$↓C$_5$↑-meander) [71]; (b) TWIQNGSTKWYQN GSTKIYT, 20–30% in β structure, H$_2$O, pH 3.25, 10˚C (C$_5$↑C$_5$↓C$_5$↑-meander) [72]; (c) RGWSLQNGKYTL NGKTMEGR, ~35% in β structure, 10%D$_2$O/H$_2$O or D$_2$O, pH 5, 0–10˚C (C$_{7eq}$↑C$_{7eq}$↓C$_{7eq}$↑-meander) [73]; (d) C$_5$↑C$_{7eq}$↑-parallel sheet, cf. the FP_i plot. The C-termini of two strands are connected by the D-prolyl-1,1-dimethyl-1,2-diaminoethane unit (diamine linker D-Pro-DADME), ~64% 'folding-core' residues (F5-V8 and R11-L14) in β structure at 10˚C, 10%D$_2$O/H$_2$O, 100 mM sodium acetate buffer, pH 3.8 [74]; (e) C$_{7eq}$↑C$_5$↑C$_{7eq}$↑-parallel sheet, cf. the FP_i plot. The C-termini of strands 1 and 2 are connected by the diamine D-Pro-DADME while the N-termini of strands 2 and 3 are connected by the diacid formed from (1*R*,2*S*)-cyclohexanedicarboxylic acid (CHDA) and Gly, 4˚C, 10%D$_2$O/H$_2$O, 2.5 mM sodium [D$_3$]acetate buffer, pH 3.8 [75]; (f) the C$_{7eq}$ strands from two C$_5$↑C$_{7eq}$↑-parallel sheets in the left-handed coils of PBP2x from *Streptococcus pneumoniae*, PDB ID 1k25; (g) the C$_5$ strands from two C$_5$↑C$_{7eq}$↑-parallel sheets in the left-handed coils of PBP2x, PDB ID 1k25.

the structure in the matrix of the ionic atmosphere, *vide infra*. In the lipid environment, the folding basin is an elevation of relative permittivity: the surface→core gradient of relative permittivity associated with the compact structure of a transmembrane protein is opposite to that associated with the soluble globules. Here the helix incorporating the 'C$_5$ strand' segment will be stable when it is exposed to the lipid bilayer while the helix incorporating the 'FP_i>>0 turn' segment will be stabilized when it is buried, e.g. by oligomerization. In this view, globular fold of a soluble protein develops to bury backbone segment whose folding potential FP_i is fine-tuned to make a well-structured fold unstable in solvent but stable in a less polar environment of protein interior; architecture and stabilization of tertiary structure are brought about by selective destabilization of secondary structure.

(iv) The effects of charge separation in the medium: Folding template and bounds of tertiary structure. The effect of mutual polarization of the protein and its environment also depends on the intrinsic charge separation in the medium which may act as *the folding template* $\mathcal{FT}$. The phospholipid bilayer can be thought of as such a folding template but beyond the obvious effect of low-permittivity environment of the lipid matrix it is not clear how its complex structural and physicochemical features would affect polarization of the polypeptide backbone. On the other hand, charge separation in the medium of cytosol or blood serum etc., i.e. in the 1:1 electrolytes (e.g. KCl, NaCl) under the standard physiological conditions, is better understood [82]. Here, the function of the folding template may be performed by the transient quasi cubic lattice of ionic atmosphere with the constant of 7 Å, the length that the Bjerrum distance and the Debye radius converge to in such solutions. The notion that the crystal-like lattice persists in dilute salt solutions was introduced by Gosh a century ago to explain nonideality of such solutions [83,84]. and was later refined by Debye and Hückel (see S1 Appendix) [48], hence the said lattice is here referred to as the Ghosh-Debye-Hückel matrix. The protein/electrolyte system is stabilized when the key surface charges are placed in the vertices of this ionic matrix; two charges e_i and e_j ought to be separated by the distance $[(7\Delta x_{ij})^2+(7\Delta y_{ij})^2+(7\Delta z_{ij})^2]^{1/2}$ (Å) to fit into the lattice, where Δx_{ij}, Δy_{ij}, Δz_{ij} are whole numbers and the sum $|\Delta x_{ij}|+|\Delta y_{ij}|+|\Delta z_{ij}|$ is odd if the e_i, e_j signs are opposite, and even if the e_i, e_j signs are same, see Fig 8. The key surface charges may be the charges carried either by the ends of helices and cross-β arrays of the H-bonded peptide bonds (capped by helix turns, reverse turns or β bulges), or by the side chains (E$^-$, K, R). Note the corollary inference that soluble globular proteins and their environment may have evolved to take advantage of nonideality of dilute 1:1 electrolyte
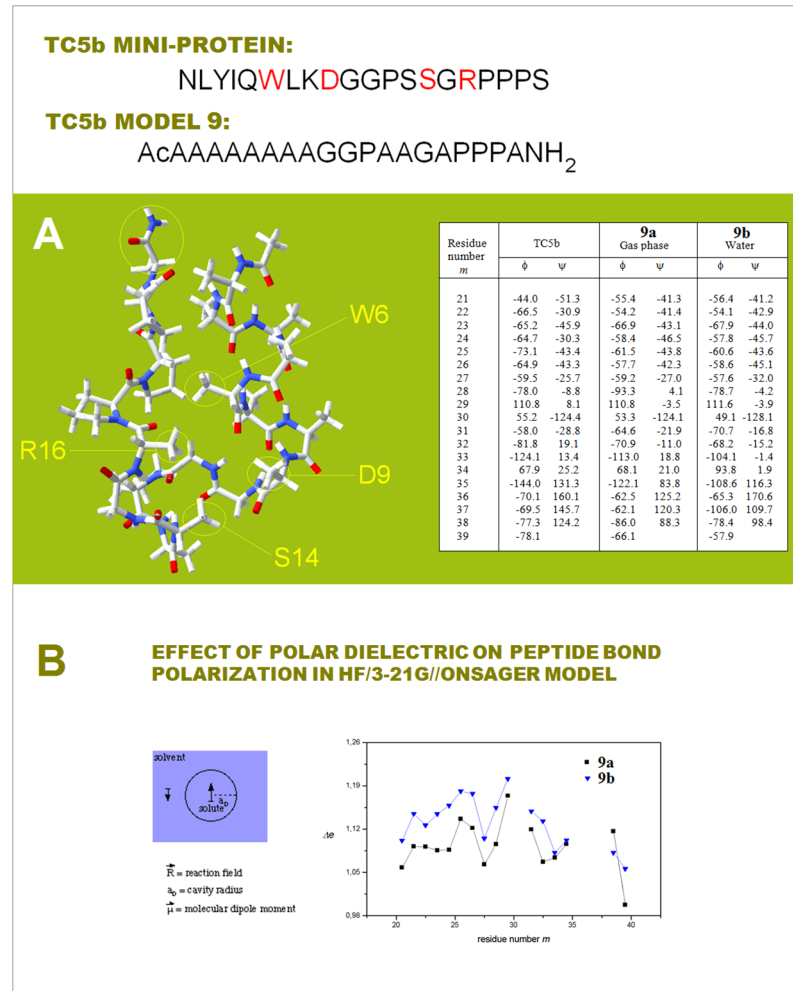
Fig 6. Effect of a polar dielectric on peptide bond polarization in a model of TC5b mini-protein. The structure of the simplified model of TC5b, **9**: AcAAAAAAAAGGPAAGAPPPA–NH$_2$, obtained by full unconstrained optimization (HF/3-21G, gas phase) of the peptide chain placed in the conformation defined by the φ and ψ angles reported for the NMR ensemble of TC5b; the final structure was re-optimized in water taken as a continuous dielectric (the Onsager model as implemented in the Gaussian suite, cf. Computational Methods), until the default convergence criteria were fully met again. The backbone torsion angles of the TC5b NMR structure PDB ID 1l2y [50] and the *ab initio* structures **9a** (in gas phase) and **9b** (in a polar dielectric) are compared in the table on the right-hand side of the panel. **(B)** Dependence of charge polarization of the secondary peptide bonds Δe (the difference (au) in H and O Mulliken populations of the *m* peptide bond [33]) on *m*—that is on bond location along the polypeptide chain in the models **9a** and **9b** of the mini-protein TC5b. The immersion of the TC5b model in a polar solvent results in the increase in Δe along the entire chain i.e. in a considerable increase in charge polarization of the polypeptide backbone.

solutions. The hypothesis that "charged cytoplasmic macromolecules are stabilized electrostatically by their ionic atmosphere" and that the geometrical dimensions of biopolyelectrolytes and their polar functionalities may be related to the physiological ionic strength, was previously advanced in the context of modelling the organization of cytoplasm [86].

    (v) Two-electron stabilizing interactions and the folding pathways of globular proteins. Taken together, the results of our modelling studies and the inferences discussed above suggest that the following factors define how an amino acid sequence 'selects' the backbone fold: (1) the intrinsic pattern of backbone polarization imprinted by the side chains' electronic
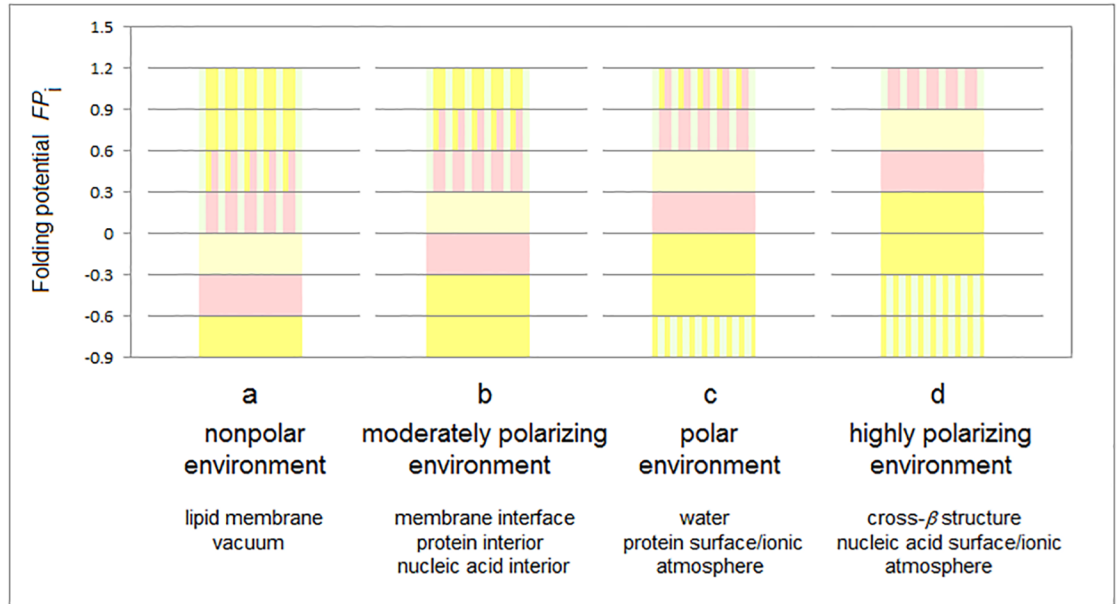
Fig 7. Folding potential, medium properties and secondary structure preferences of the polypeptide backbone.
(a) The FP_i values that ensure stability of the periodic secondary structure in a non-polar environment such as the lipid matrix of the bilayer membrane or vacuum: the optimal FP_i range for the α-helix' is –0.6-–0.3 (color-coding as in Fig 1) and the optimal FP_i ranges for β structure is <–0.6 (C_5 strand) and –0.3–0 (C_{7eq} strand). The less polarized segments are malleable in a non-polar aprotic medium and may adopt helical (3_1-helix, PP_{II}-helix, α*-helix) folds while the least polarized segments of the polypeptide backbone, e.g. a sequence of consecutive 'FP_i>>0 turns' (Fig 5), may adopt the extended (C_5* strand) folds depending on molecular embedding. **(b)** The FP_i values that ensure stability of the periodic secondary structure in a moderately polarizing environment such as the bilayer membrane interface, the interior of a soluble protein globule or the interior of the DNA duplex: the optimal FP_i range for the α-helix' is –0.3–0 and the optimal FP_i ranges for β structure is <–0.3 (C_5 strand) and 0–0.3 (C_{7eq} strand). **(c)** The FP_i values that ensure stability of the periodic secondary structure in a polar medium such as the physiological 1:1 electrolyte solution: the range of the optimal FP_i values for the α-helix is now 0–0.3 while the somewhat less and more polarized segments are likely to form β-sheets. The most polarized segments are now likely to form 'FP_i<<0 turns' or PP_{II}-helix. The sequence of consecutive 'FP_i>>0 turns' forms a random coil in an aqueous buffer unless it is stabilized by molecular embedding in helical (3_1-helix, PP_{II}-helix, α*-helix) or extended (C_5* strand) folds. **(d)** The FP_i values that ensure stability of the periodic secondary structure in the hypothetical highly polarizing environment such as the pre-organized ionic grid e.g. on the surface of a DNA or RNA strand (the sequence of consecutive 'FP_i>>0 turns' is likely to form here an α*-helix), or the microenvironment of the extended β structure of β solenoid or amyloid filament, *vide infra*.

effects (FP), (2) the emerging shell→core elevation or depression of relative permittivity and the placement and sequestration of side chains' charges (FB), and (3) the constraints of an ionic or lipoid matrix (FT). Each factor's effect varies along the folding pathway in a manner dependent on the contributions of the other factors, and each factor impacts the conformation of the protein by controlling, directly or indirectly, electronic configuration and bonding interactions of the peptide amide linkages. Thus, we propose that the pathway of folding of a globular protein comprises a sequence of conformational transitions driven by changes in the free energy of the polypeptide backbone which, to a considerable degree, are determined by the two-electron stabilizing interactions such as the generalized anomeric effect, homohyperconjugation of peptide linkages and covalent contributions to backbone-backbone H-bonding. The well-studied folding of the small soluble helix-bundle domains presents a system which seems to behave in this way: the 'helix' propensity of the domain, defined by the folding potential FP_i, appears to have considerable impact on the folding rate constant k_f^{H2O}, Fig 9 [87–100]. This is consistent with the transition state ensemble having an approximately native topology but no fixed shell→core gradient of relative permittivity and no effective interaction with the
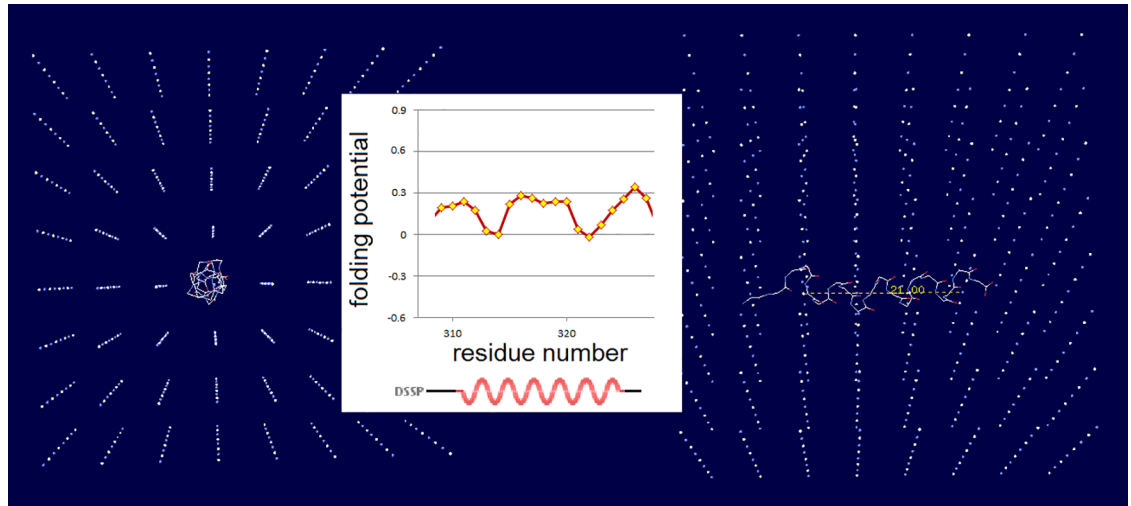
Fig 8. Folding potential, folding template and three-dimensional structure of soluble globular proteins. The insert shows the plot of the folding potential FP_i for the segment of the polypeptide backbone which has high helical propensity in the aqueous environment: the 14-residue site which triggers coiled-coil formation in cortexillin I [85]. In the physiological 1:1 electrolyte solution, this segment is stabilized by the mutual polarization of the α-helix and the transient ionic matrix with the lattice constant of 7 Å (the Ghosh-Debye-Hückel matrix, see the text and S1 Appendix). The effect of polarization is maximized when the helix termini replace the corresponding salt ions in the vertices of the lattice which are separated by the distances $[(7\Delta x_{ij})^2+(7\Delta y_{ij})^2+(7\Delta z_{ij})^2]^{1/2}$ (Å) where Δx_{ij}, Δy_{ij}, Δz_{ij} are whole numbers and the sum $|\Delta x_{ij}|+|\Delta y_{ij}|+|\Delta z_{ij}|$ is odd. Thus the 'allowed' α-helix is a vector whose length is defined by the $|\Delta x_{ij}|,|\Delta y_{ij}|,|\Delta z_{ij}|$ combinations equal to: {1,0,0}/7 Å, {1,1,1}/12 Å, {2,1,0}/15.6 Å etc. The length of the α-helix in the diagram is 21 Å which fulfils the above condition when the helix fits into the matrix along the grid line ($|\Delta x_{ij}|,|\Delta y_{ij}|,|\Delta z_{ij}|$ = {3,0,0}, $|\tau_{min}|$ = 0˚), as shown here in both projections, or along the diagonal of the 2×2 segment of 4 unit cells ($|\Delta x_{ij}|,|\Delta y_{ij}|,|\Delta z_{ij}|$ = {2,2,1}, $|\tau_{min}|$ = 48˚) (where $|\tau_{min}|$ is the smallest vector/grid-line angle).

transient lattice of the ionic atmosphere, so that the helices which incorporate the 'strand' and 'turn' segments are not fully stabilized in the transition state. Such a stabilization is eventually achieved in the native state and the dependence of the free energy of folding $\Delta G_{U\text{-}F}^{H2O}$ on 'helix' propensity is obscured.

b. Mechanism and principles of encoding the 3D structure of proteins: Explanatory and predictive power of the PMO model

(i) Stability of secondary structure as quadratic or quartic function of the folding potential *FP*. To probe the dependence of the free energy $\Delta G_{coil\rightarrow helix}$ and $\Delta G_{coil\rightarrow sheet}$ on the electronic configuration of the polypeptide backbone, we examine a range of phenomena including single-site mutagenesis, amide-ester substitution, amyloidogenic propensity, and molecular recognition of PDZ domains. In Fig 10, thermodynamic secondary structure propensities [101–114] are plotted against the calculated tensors $\sigma(C^\alpha)^{Xaa}$ (see the Computational Methods); these plots are consistent with the expected quadratic and quartic dependence for the α and β structure respectively.

In Fig 11, we test the average value of the FP_i function, $\overline{FP_i}$, as a measure of the conformational and H-bonding propensity of a segment of the polypeptide chain. Several lines of evidence confirm that the $\overline{FP_i}$ values may indeed carry such information: (1) the plots of average temperature factors B_i indicate one minimum of backbone mobility with respect to $\overline{FP_i}$ in α-helices, and two such minima in the strands of β structure, Fig 11A(a) and 11A(b) [115]; (2) the $\Delta(\Delta G_f)$ data on the amide-to-ester substitutions suggest that the energy of backbone-backbone H-bonding in the β-sheet of Pin1 WW domain has two minima with respect to $\overline{FP_i}$ at the
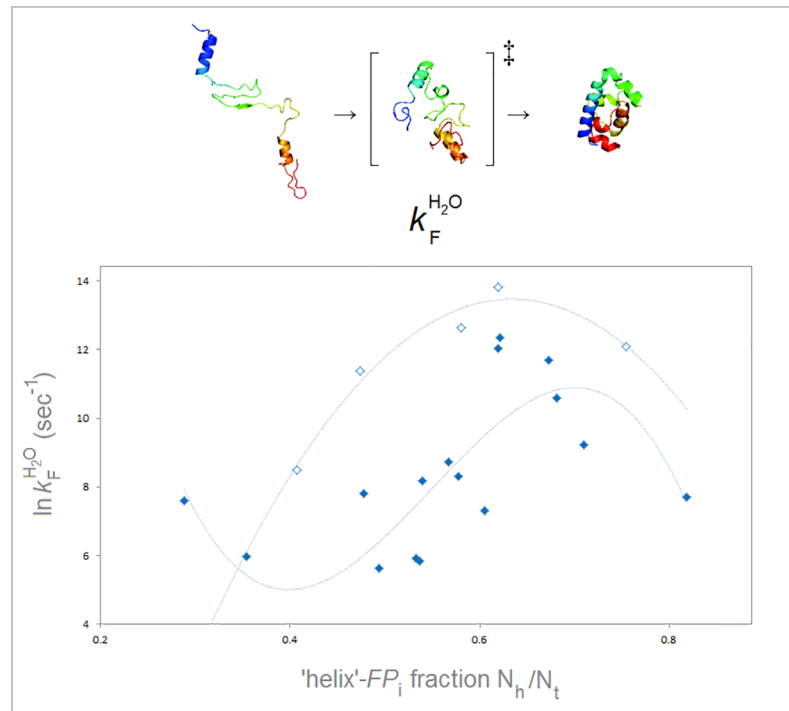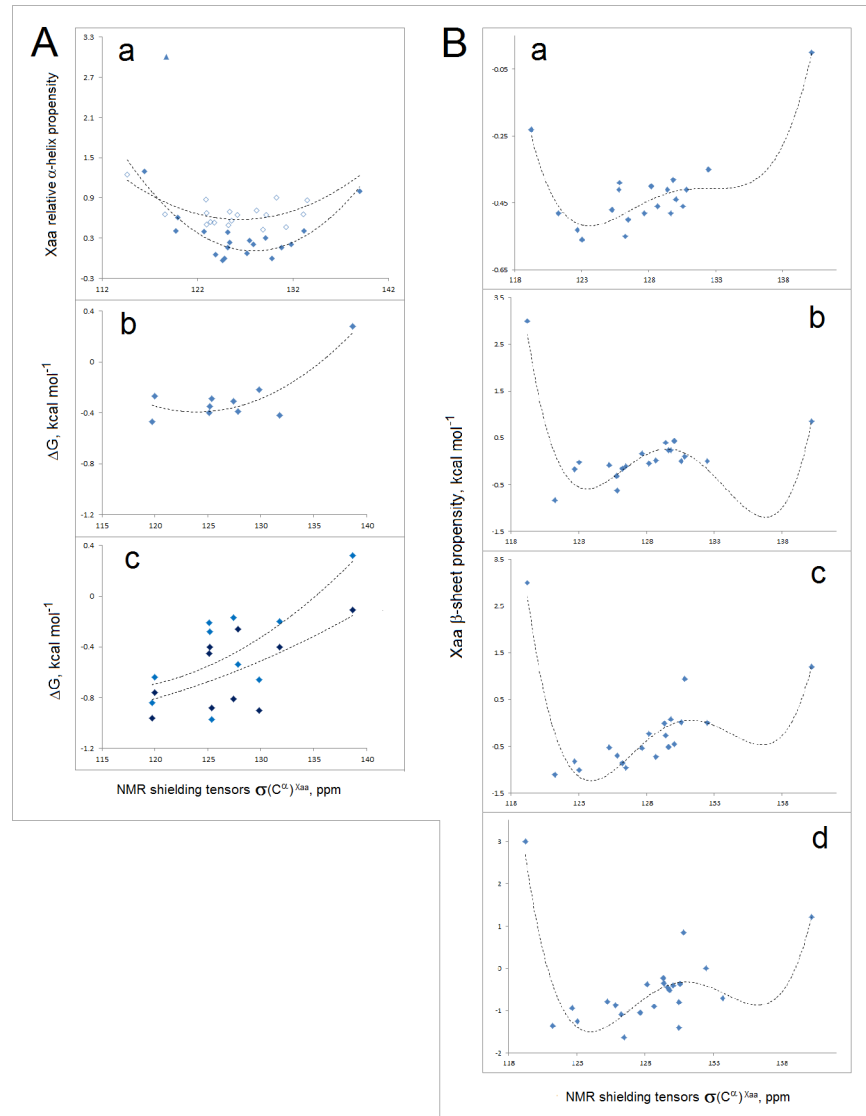
Fig 9. Electronic configuration of the polypeptide backbone and rate of folding of helix-bundle proteins. The folding rate constants $\ln k_f$ (sec^{-1}) vs. N_h/N_t ('helix'- FP_i fraction) where N_h is the number of residues with the 'helix' FP_i: 0 ± 0.05–0.3 ± 0.05, cf. Fig 7(c), and N_t is the total number of residues in the helix-bundle domains, counted from the N-terminal residue of the first α-helix to the C-terminal residue of the last α-helix as defined by the DSSP protocol implemented in the RCSB PDB database. The present set includes the data for 16 small proteins with the natural, wild-type sequences and for 5 domains modified or engineered for fast folding [87–100] $N_t \leq$ ~80 aa, PDB ID's: 1ayi, 1ba5, 1fex, 1imp, 1mbk, 1prb, 1ss1, 1st7, 1uzc, 1yrf, 2a3d, 2abd, 2jws, 2jwt, 2no8, 2wqg, 3kz3: ♦ wild-type domains; ◊ the domains modified/engineered for fast folding.

site of substitution, Fig 11B [116,117]; and (3) amyloidogenic propensity of linear hexapeptides appears to display two maxima with respect to the peptide $\overline{FP_i}$ [118], Fig 11C.

The characterization of the conformational and H-bonding propensity of a segment of the polypeptide backbone in terms of $\overline{FP_i}$ may also be valid in more complex systems. The canonical binding of oligopeptides by the PDZ domains [119] involves extension of the domain's β structure: the oligopeptide is inserted into the binding pocket as the edge strand of the antiparallel β sheet. The plots of binding affinity ΔG_b [120–123] against the average FP_i value of the peptide ligand, $\overline{FP_i}$ (peptide), seem to confirm the expected quartic dependence of ΔG_b with respect to $\overline{FP_i}$, see Fig 12.

Lastly, the $\Delta(\Delta G_f)$ differences in stability of the large-to-small hydrophobic variants—used to estimate the free energy of hydrophobic interactions [124], appear to stem in good part from the changes in the conformational and H-bonding propensity of the main chain as well, see Fig 13. The Xaa→Ala mutation (Xaa = F, I, L, M, T, V, W, Y) may destabilize the native state by changing, among others, backbone's folding potential FP_i. The deleterious effect of such a change is expected to be particularly significant when the mutation occurs in the region of high congruence of the folding potential, environment and secondary structure, i.e. in the well-ordered segment that anchors the fold. Thus, the difference $\Delta(\Delta G_f)$ in stability of the Xaa→Ala hydrophobic variants should have one minimum with respect to FP_i in α-helices but two such minima in β-sheet strands. The available data seem to support this notion: the plots

Fig 10. Electronic configuration of the polypeptide backbone and secondary structure propensity. (A) Experimental α-helix propensities: (a) The averaged relative α-helix propensity data obtained in the site-directed mutagenesis studies of both peptides and proteins, adjusted so that $\Delta(\Delta G_f) = 0$ for Ala and $\Delta(\Delta G_f) = 1$ for Gly [101–108], vs. the NMR shielding tensors $\sigma(C^\alpha)^{Xaa}$ (3_{10}-helix AcG(Xaa)GGGNH$_2$; GIAO//B3LYP/D95**, cf. Computational Methods and S1 Table): ♦ glycine and amino acids whose C^β and C^γ are the methyl, methylene or methine groups, $r^2 = 0.83$; ▲ proline; ◊ any other amino acids including three highly fluorinated amino acids, $r^2 = 0.52$ [107]; trendlines obtained by fitting 2nd order polynomial functions; (b) The Lifson-Roig propagation free energies for the amino acids whose C^β and C^γ are the methyl, methylene or methine groups, in 88% methanol-water [109]; (c) The Lifson-Roig propagation free energies for the same set of amino acids in 40% (cyan) and 90% (navy) trifluoroethanol-water [109]. The propensities are determined at the sites in the helices interior. **(B) Experimental β-sheet propensities** from site-directed mutagenesis (kcal mol^{-1}, $\Delta(\Delta G_f) = 0$ for Gly in (D) and $\Delta(\Delta G_f) = 0$ for Ala in (E), (F) and (G)) vs. calculated NMR shielding tensors $\sigma(C^\alpha)^{Xaa}$ (AcGGGGGXaaNHMe in β-hairpin (Ib turn); GIAO//B3LYP/D95**, cf. Computational Methods and S1 Table): (a) zinc-finger β-hairpin, site 3, $r^2 = 0.89$ (edge strand, the guest site is not H-bonded) [110]; (b) Ig binding B1 domain of streptococcal protein G, $r^2 = 0.83$ (variant E42A/D46A/T53A, site 44, edge strand, the guest site is H-bonded) [111]; (c) Ig binding B1 domain of streptococcal protein G, $r^2 = 0.84$ (variant I6A/T44A/T51S/T55/S, site 53, central strand) [112]; (d) Ig binding B1 domain of streptococcal protein G, $r^2 = 0.76$ (I6A/T44A, site 53, central strand) [113,114]. $\Delta(\Delta G_f)$ for Pro in (b), (c) and (d) set at the minimum value of 3 kcal mol^{-1} [112]; trendlines obtained by fitting 4th order polynomial functions.

https://doi.org/10.1371/journal.pone.0180905.g010
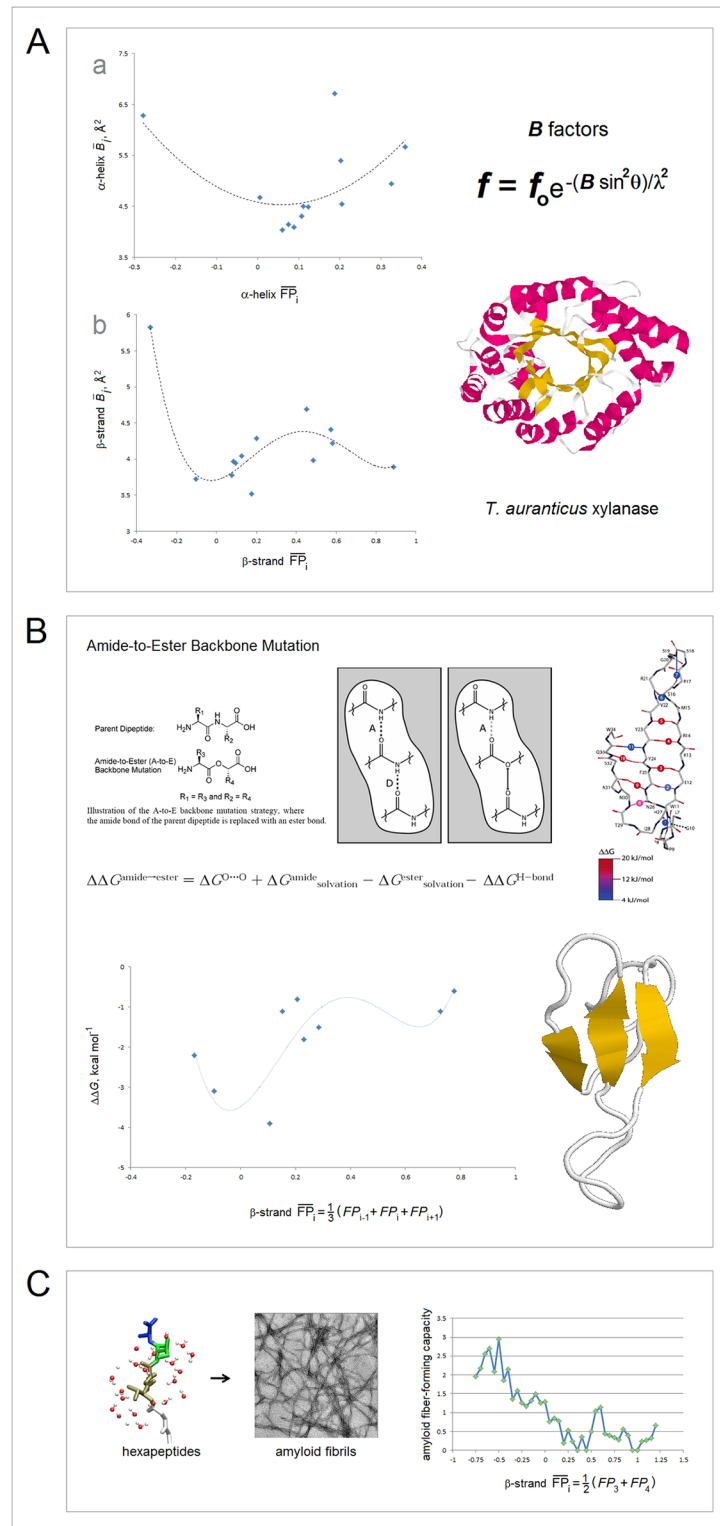
Fig 11. $\overline{FP_i}$ as a measure of conformational and H-bonding propensity of the polypeptide backbone.
(A) $\overline{FP_i}$ as a probe of backbone mobility: (a) The mean temperature factors B_i of the backbone N atoms in
α-helices vs. mean FP_i of those helices, $\overline{FP_i}$ (α), in the xylanase from *Thermoascus auranticus*, PDB ID 1i1wA
[115]. Helical residues are assigned according to the Swiss-PDBViewer: helix **A** 6–12, **B** 24–27, **C** 32–38, **D**
51-54, **E** 64–76, **F** 93–96, **G** 101–117, **H** 143–147, **I** 151–163, **J** 182–197, **K** 215–227, **L** 245–257, **M** 292-301;

(b) The mean temperature factors B_i of the backbone N atoms in β strands vs. mean FP_i of those strands, $\overline{FP_i}$ (β). The strand residues are assigned according to the Swiss-PDBViewer and DSSP protocol implemented in the RCSB PDB database: **N** 17–22, **O** 41–46, **P** 79–81, **Q** 124–127, **R** 132–134, **S** 138–140, **T** 168–173, **U** 202–206, **V** 208–210, **W** 232–236, **X** 239–242, **Y** 264–266, **Z** 279–281. The trendlines are obtained by fitting 2nd and 4th order polynomial functions. **(B)** $\overline{FP_i}$ **and the energy of backbone-backbone H-bonding**. $\Delta(\Delta G_f)$ vs. FP_i for the single-site amide-to-ester X(i)ξ substitutions ($\Delta(\Delta G_f) = \Delta G_{f,WT} - \Delta G_{f,X(i)ξ}$) in Pin1 WW domain. The data shown for the mutants in which the perturbed amide donates, but does not accept, a hydrogen bond (thermal, GdnHCl, pH 7.0; PDB ID 2kcf) [116,117]. The trendline is obtained by fitting 4th order polynomial function. **(C)** $\overline{FP_i}$ **and the amyloid fibril-forming capacity of oligopeptides**. The data for the total of 942 unique hexapeptide structures are taken from the WALTZ-DB database of amyloid forming peptides [118]; excluding the Pro-containing peptides, the sample comprises 240 amyloidogenic and 702 non-amyloidogenic hexapeptides. The mean FP_i of each hexapeptide, $\overline{FP_i}$ (peptide), is defined as the mean of FP_i of the two central residues. The amyloid fibril-forming capacity of the hexapeptides with $\overline{FP_i}$ within a specific 0.05 range (–0.750±0.025 etc.) is defined as the frequency of the amyloidogenic peptides within this $\overline{FP_i}$ range in the entire amyloidogenic sample (240 entries), normalized by the frequency of both amyloidogenic and non-amyloidogenic peptides within the same $\overline{FP_i}$ range in the total hexapeptide sample (942 entries).

in Fig 13 show that $\Delta(\Delta G_f)$ has one minimum with respect to $\overline{FP_i}$ at the mutation site in α-helices [87,125–128] (Fig 13A), and two such minima in β-sheets [125,129–131] (Fig 13B).

(ii) Stability of secondary structure as a function of environment: Continuous dielectric. The plots of Lifson-Roig propagation free energies obtained in the 88%methanol-water and 40% and 90% trifluoroethanol-water mixtures, Fig 10A(b) and 10(c), when compared to the plots of α-helix propensities in water in Fig 10A(a), confirm that the expected shift of secondary structure propensity, cf. Fig 7, indeed takes place. These data are consistent with the data for other nonpolar environments e.g. lipid micelles and vesicles [132], and the gas phase, where Val and Ile are better helix-formers than Ala [133].

(iii) Stability of secondary structure as a function of microenvironment: Extended β structure. Polarization of the polypeptide backbone, and therefore secondary structure propensity, will be affected by the interactions with charges and partial charges within molecular embedding. Indeed, H-bonding interactions of the peptide amide bonds with ions e.g. within the 'capping' motifs or within the pre-organized ionic grid/ionic atmosphere of the DNA duplex surface [134,135], are expected to be highly polarizing. However, given the cooperativity of backbone-backbone H-bonding [136], the cross-β H-bonded arrays of extended β-structure elements, including asparagine and glutamine ladders, may also create a highly polarizing environment; the extraordinary dielectric properties of low-MW secondary amides are attributed to the formation of such extended arrays in bulk liquids [137]. Investigations of the amide I region in FTIR indicate that the microenvironment of amyloid fibrils does strengthen backbone H-bonding [138,139]; the maxima of amyloidogenic propensity, and especially the location of the 'helical' region (the minimum) at the $\overline{FP_i}$ (peptide) values between 0.25 and 0.50, Fig 11C, suggest that the amyloid's microenvironment is indeed highly polarizing, cf. Fig 7(d).

The canonical binding of oligopeptides by the PDZ domains [119] previously examined in Fig 12, seems to be controlled by this effect. The data in Fig 12 are insufficient to fully define the minima in all the plots of the binding affinity ΔG_b vs. $\overline{FP_i}$ (peptide) but the positions of the ΔG_b maximum are accurate enough to say that these maxima shift between the $\overline{FP_i}$ (peptide) values of ca. –0.25 and 0.25. We expect, cf. Fig 7, that the shift occurs because the binding pockets of the PDZ domains differ in the capacity to polarize the peptide ligands; the correlation of the $\overline{FP_i}$ (peptide) values at the ΔG_b maximum with the FP_i of the PDZ cross-β structure, see Fig 14 and S3 Appendix, validates this notion.

(iv) Stability of tertiary structure as a function of the folding potential *FP* and the folding basin *FB*. The architecture and stabilization of tertiary structure are brought about, we
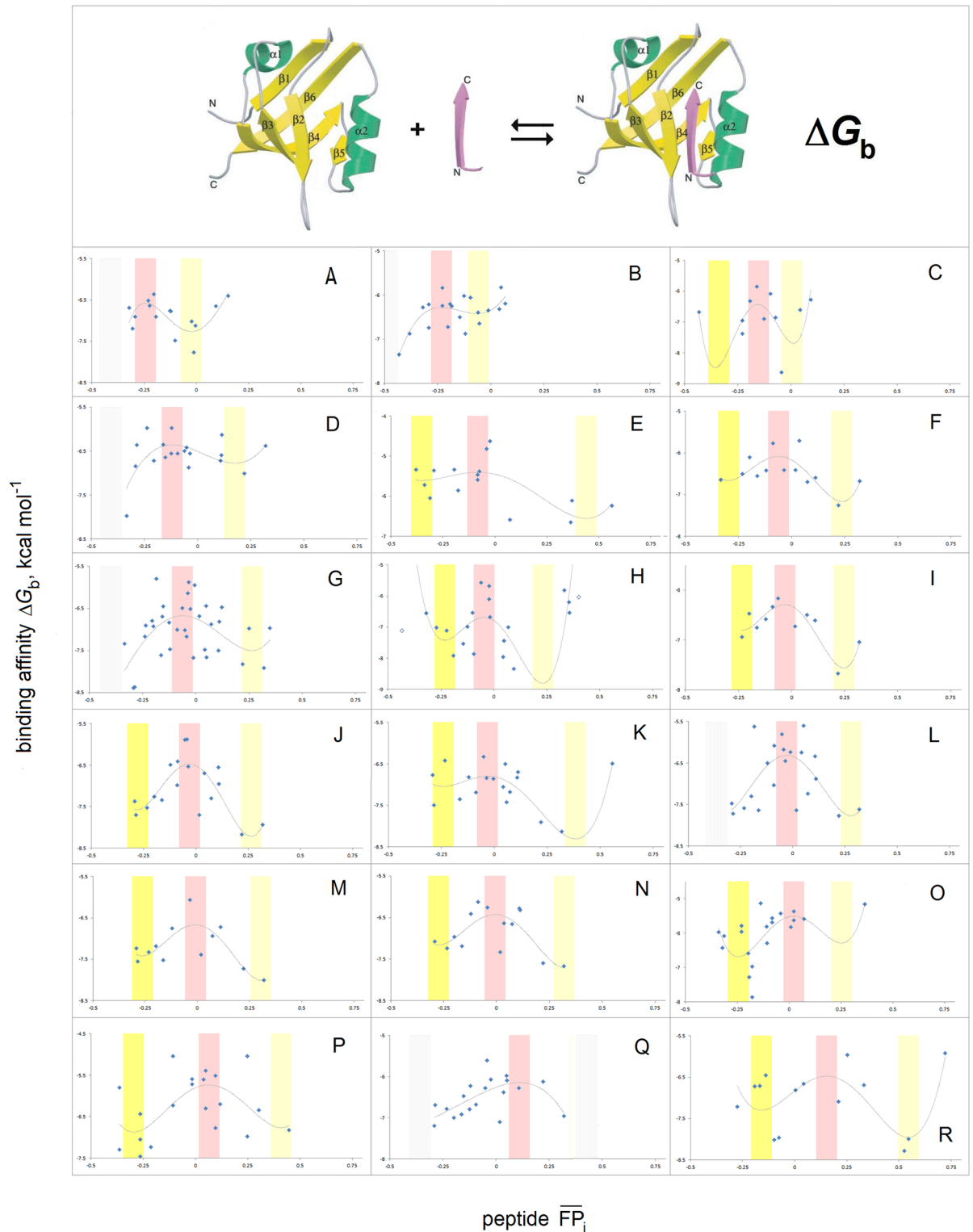
Fig 12. Electronic configuration of the polypeptide backbone and canonical peptide recognition by the PDZ domains.
Binding affinities ΔG_b of the PDZ domains vs. mean FP_i of the peptide ligands $\overline{FP_i}$ (ΔG_b in kcal mol^{-1}, the width of the FP_i averaging window is given in the brackets). The data from S2 and S3 Tables in ref. [120] unless indicated otherwise, for 18 out of 33 PDZ domains identified in the literature to bind at least a dozen or so peptides with K_d<100 μmol (no discernible dependence of ΔG_b on $\overline{FP_i}$ in 3 cases, r^2<0.25 in 11 cases, and in 1 case the fit depends on an outlier) [120–123]: **(A)** MAGI2 PDZ2, $r^2 = 0.52$ (5); **(B)** PTP-BL, $r^2 = 0.52$ (5); **(C)** MAGI1 PDZ6, $r^2 = 0.32$ (8); **(D)** Lin7C, $r^2 = 0.38$ (8); **(E)** AF6, $r^2 = 0.42$ (5) [121]; **(F)** SAP97 PDZ1,

have proposed in section **a**.(iii), by selective destabilization of secondary structure that aims to take advantage of the permittivity gradient of the protein's folding basin *FB*. For instance, to form a compact bundle of α-helices in an aqueous buffer, protein sequence ought to comprise some backbone segments which have 'helical' $\Delta FP_{i-1 \to i+1}$ profiles and optimal 'helical' $\overline{FP_i}$ values (0–0.3 in water, cf. Figs 5 and 7) and will fold into solvent-exposed helices. At the same time, however, such a sequence also ought to incorporate backbone segments which have 'helical' $\Delta FP_{i-1 \to i+1}$ profiles but 'non-helical' *depressed* $\overline{FP_i}$ values (≤ 0, cf. Fig 7) and will fold into helices only in a low-permittivity environment e.g. when they are buried in the protein interior. The presence of those 'destabilized-helix' segments is necessary to organize and stabilize the tertiary structure of such a bundle.

Thus, as the complexity of the native fold and the order of oligomerization of *soluble* proteins increase (the folding basin *FB* increases in size), backbone's folding potential *FP* is expected to become more negative so that the FP_i averages for the major secondary structure elements are expected to shift to lower values. In contrast, cf. section **a**.(iii) and Fig 7, in the case of the *integral membrane* proteins, the FP_i averages of those elements are expected to shift to higher values as the complexity of structure and the size of the protein's folding basin *FB* increase. A survey of the representative soluble and membrane proteins seems to validate this concept, see the FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots in Fig 15.

In principle, protein-ligand interaction may also be described in terms of reorganization of the folding basin *FB*. The complex is likely to change backbone conformation and relative permittivity of binding site's environment, and the $\overline{FP_i}$ s of the segments directly involved in the binding may deviate from the optimal values to accommodate such change. The $\overline{FP_i}$ variation in chemokines, chromodomains, and bZIP proteins, is consistent with this notion, see Fig 16. The chemokine and CHROMO domains have similar β1-β2-β3-α1 folds but the roles of the N-terminal and C-terminal segments are reversed: chemokines use the helix to bind to GAGs [140], while CHROMO domains use the β1-β2 hairpin to bind to N-Me$_3^+$ Lys of histones [141]. Accordingly, chemokines tend to have low-$\overline{FP_i}$ helices which will be stabilized in a low-permittivity environment and stable $C_5\uparrow C_5\downarrow$ hairpins, Fig 16A(a), while chromodomains tend to have stable solvent-exposed helices and low-$\overline{FP_i}$ $C_{7eq}\uparrow C_{7eq}\downarrow$ hairpins which will be stabilized in a low-permittivity environment, Fig 16A(b).

The bZIP proteins are long helical dimers which bind to the major groove of DNA duplex: the basic N-terminal segment interacts with DNA while the C-terminal leucine zipper domain triggers formation of the coiled-coil structure [85] and stabilizes the dimer. The FP_i plots for a multiple alignment of bZIP proteins show how each function is related to a distinct $\overline{FP_i}$ value, see Fig 16B (the assignments based on Fig 7): (1) the N-terminal segment has the 'C_{7eq} strand'-FP_i i.e. the FP_i of a helix that is stable in a highly polarizing environment rather than in water; (2) the central segment has the 'C_5 strand'-FP_i i.e. the FP_i of a helix that is stable in a nonpolar environment rather than in water and therefore drives the aggregation of the coiled coil; and (3) the C-terminal segment has the 'helix'-FP_i i.e. it is expected to form a stable in water α-helix which triggers formation of the coiled-coil structure [85].

(v) Stability of tertiary structure as a function of the folding template *FT*. As was argued in section **a**.(iv), mutual polarization of the protein solute and the transient cubic lattice
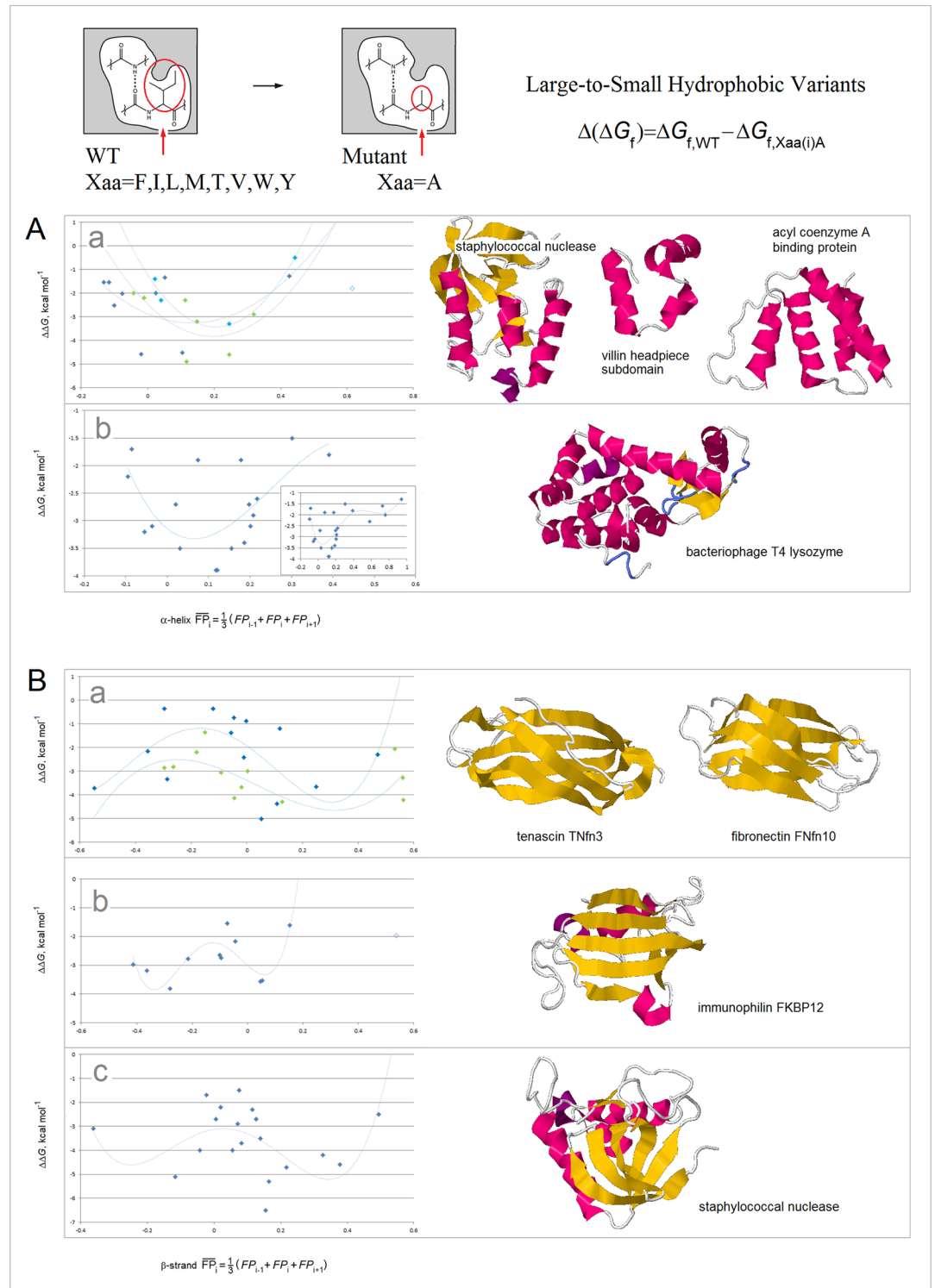
Fig 13. Electronic configuration of the polypeptide backbone and stability of large-to-small hydrophobic variants. The figure presents the $\Delta(\Delta G_f)$ vs. $\overline{FP}_i$ plots for the single-site Xaa(i)Ala mutations in helices and sheets ($\Delta(\Delta G_f) = \Delta G_{f,WT} - \Delta G_{f,Xaa(i)Ala}$, $\overline{FP}_i = (FP_{i-1} + FP_i + FP_{i+1})$); Xaa = F, I, L, M, T, V, W, Y). The 'helix' and 'strand' residues are assigned based on the DSSP protocol implemented in the RCSB PDB database; the trendlines are obtained by fitting polynomial functions [125–131]: **(A)** The $\Delta(\Delta G_f)$ data for the α-helices in: (a) staphylococcal nuclease (GdnHCl, pH 7.0 [125], PDB ID 1nuc), villin headpiece subdomain (thermal, pH 7.0 [126], PDB ID 1yri, the data set of the

highest $\overline{FP}_i$ (shown) is not included in the calculation of the trendline), and acyl coenzyme A binding protein (GdnHCl, pH 5.3 [87], PDB ID 2abd); (b) bacteriophage T4 lysozyme (thermal, pH 3.0 [127,128], PDB ID 2lzm, the complete scattergram, including the four data sets of the highest $\overline{FP}_i$, is shown in the insert). **(B)** The $\Delta(\Delta G_f)$ data for the β-sheet strands in: (a) fibronectin type III domains of human tenascin TNfn3 (3rd module, PDB ID 1ten) and fibronectin FNfn10 (10th module, PDB ID 1fnf) (thermal, urea, GdnHSCN, pH 5.0) [129,130]; (b) immunophilin FKBP12 (urea, pH 7.5 [131], PDB ID 2ppn, the data set of the highest $\overline{FP}$ (shown) is not included in the calculation of the trendline); (c) staphylococcal nuclease (GdnHCl, pH 7.0 [125], PDB ID 1nuc, the data set of the highest $\overline{FP}_i$ (not shown) is not included in the calculation of the trendline).

of the ionic atmosphere in the physiological 1:1 electrolyte solution may stabilize the system if the key surface charges of the protein replace the corresponding salt ions in the vertices of the ionic matrix. For instance, in the case of globins, the protein/electrolyte system is likely to be stabilized when the termini of the α-helices are placed in those vertices, while in the case of the congeners of *Borrelia* spirochete antigen (outer surface protein A) the protein/electrolyte system is likely to be stabilized when the alkyl ammonium ions of the lysine side chains replace the salt cations in the vertices of the ionic matrix. The expected arrangement is indeed found,
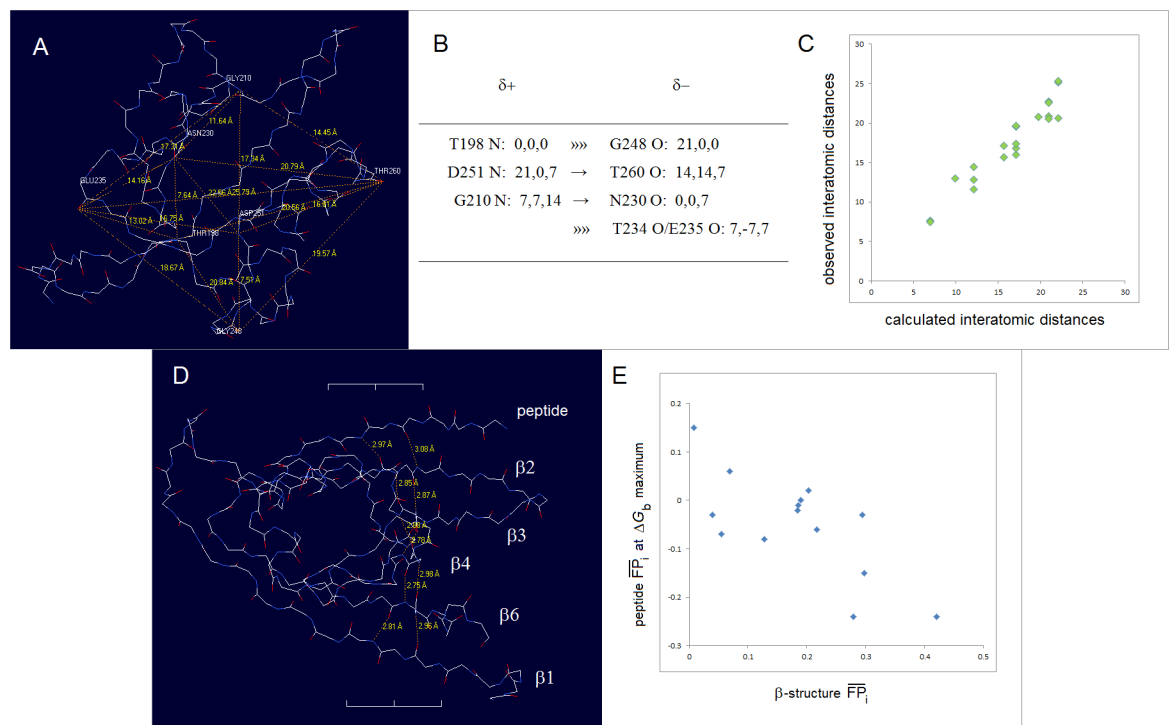
Fig 14. Electronic configuration of the polypeptide backbone and canonical peptide recognition by the PDZ domains.
(A) Folding template *FT* and the 3D structure of the PDZ domains: the putative 'key' surface charges of the PDZ fold are the termini of the α1 helix-[CO_2^- loop] and the α2 helix arrays (→), and the cross-β arrays (»») capped by a reverse turn (G248 (C' = O)) and a bulge (T234 (C' = O)/E235 (C' = O)). The structure in the diagram is the second PDZ domain of syntenin, PDB ID 1r6j. **(B)** The projected fit of the putative key surface charges δ^+ and δ^- into the Ghosh-Debye-Hückel matrix. **(C)** The predicted by the folding-template model and the observed interatomic distances (Å) between the key surface charges δ^+ and δ^- (the average distances to the T234 O/E235 O atoms in the bulge are used). **(D)** Two cross-β {β1-β6-β4-β3-β2} arrays of peptide bonds that anchor the oligopeptide ligand via backbone-backbone H-bonds to N-H of the residue $i = -2$ and to C' = O of the residue $i = -4$. **(E)** The oligopeptide $\overline{FP}_i$ values at the ΔG_b maximum (determined from Fig 12 for the 14 entries with β sheet register confirmed by the X-ray or NMR structure determination) plotted against $\overline{FP}_i$ (PDZ/sheet) i.e. the average FP_i values of the 15 residues involved in the two cross-β arrays which are shown in the panel D. For a detailed account of the model of peptide recognition by the PDZ domains see S3 Appendix.
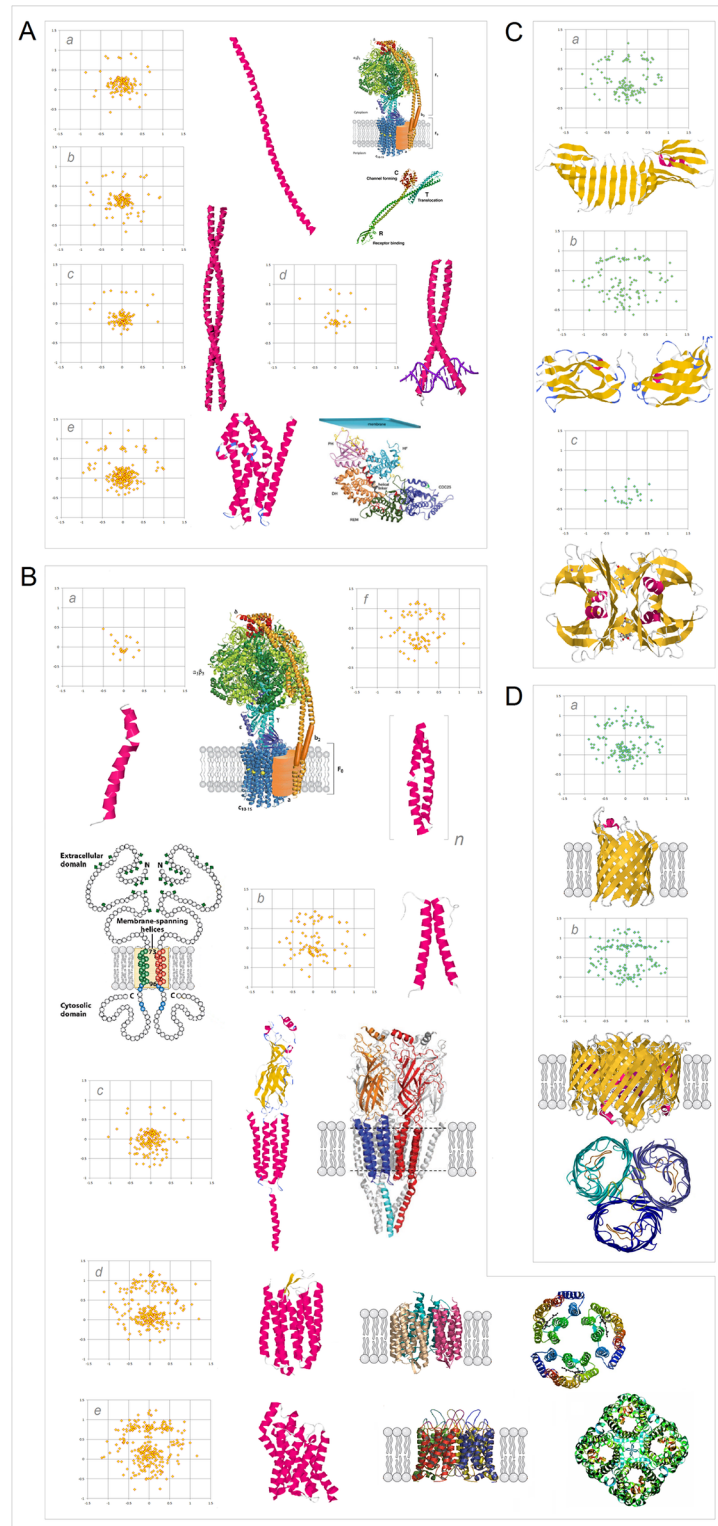
Fig 15. Electronic configuration of the polypeptide backbone versus complexity of tertiary structure and order of oligomerization in soluble and integral membrane proteins. The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots, cf. Fig 5, may characterize the relationship between distribution of backbone density and organization of globular structure. For α-helices, narrow distribution of the folding potential values and the slope values, $\Delta FP_{i-1 \to i+1} \sim 0$, generates compact clusters of the data points. As the complexity of α structure increases and the exposure to

the medium decreases, this clusters shift from $FP_i > 0$ to $FP_i < 0$ region in the case of soluble proteins, see the panels in **(A)**, and in the opposite direction, from $FP_i < 0$ to $FP_i > 0$ region in the case of integral membrane proteins, see the panels in **(B)**. The same trends are discernible in the FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots for the antiparallel β sheets that assemble 'C_5 strands' or 'C_{7eq} strands' and '$FP_i \gg 0$ turns', panels **(C)** and **(D)**, even though large differences in the folding potential values and wide distribution of the slope values, from $\Delta FP_{i-1 \to i+1} \sim -1$ to $\Delta FP_{i-1 \to i+1} \sim 1$, generate in this case circular distribution of data points. **(A)** (a) The peripheral stalk helix of F_0F_1 ATP synthase from *E. coli* (b_2 subunit in the top diagram in the right-hand panel), UniProt # P0ABA0: residues A32-K122 (the hinge region, the dimerization region, and the C-terminal δ-domain-binding region); (b) One of the two 160Å-long helices of colicin Ia (residues R351-K470) that span the periplasmic space, linking the receptor-binding domain to the other domains, PDB ID 1cii; (c) The parallel α-helical coiled coil cortexillin I, PDB ID 1d7m; (d) The Leu-zipper segment of the GCN4 bZIP protein PDB ID 2dgc: the 'helix' cluster is shifted to the $FP_i \sim 0$ region; (e) The pattern of the multi-helix bundle which functions as an enzyme within a heterooligomeric complex: the 'helix' cluster is partly shifted into the $FP_i < 0$ region. The structure shown is the catalytic domain of the guanine nucleotide exchange factor (Ddl homology (DH) domain) of human PAK-interacting exchange protein, PDB ID 1by1. **(B)** (a) Transmembrane segment of the peripheral stalk of F_0F_1 ATP synthase from *E. coli* (b_2 subunit in the diagram in the centre, color-coded orange), UniProt # P0ABA0: residues A11-A31; (b) Transmembrane α-helices of human glycophorins A, B, C and E: UniProt #'s P02724, P06028, P15421, and P04921. The glycophorin helices apparently are brought together in the membrane by the dimerization of the extra-membrane domains; (c) The transmembrane α subunit of the membrane associated acetylcholine receptor from *Torpedo marmorata*, PDB ID 2bg9; the five transmembrane subunits of this receptor do not form a tight oligomer structure and are held in place by the extra-membrane subunits; (d) Bacteriorhodopsin, a membrane protein (light-driven proton pump) from *Halobacterium salinarum*; the biological assembly is the homotrimer, PDB ID 1fbb; (e) Glycerol facilitator (GlpF), a membrane channel protein of the aquaporin family; the biological assembly is a homotetramer, PDB ID 1fx8; (f) Subunit *c* of F_0F_1 ATP synthase from *E. coli* (color-coded blue in the diagram in the centre): the biological assembly is a homodecamer, PDB ID 1ijp. **(C)** (a) The single-sheet protein, the central β-sheet (residues A81-D205) of the *Borrelia burgdorferi* spirochete antigen, outer surface protein A (OspA), PDB ID 2g8c; (b) The β sandwich C-terminal domain of the α-amylase from *Geobacillus stearothermophilus*, PDB ID 1qho; (c) The strands βF(H88-A97) and βH(S115-T123) of the homotetramer of transthyretin, buried in the tetramer interior, PDB ID 5l4j. **(D)** (a) 16-stranded β barrel, the monomeric integral outer-membrane porin OmpF from *E. coli*, PDB ID 1opr; (b) 16-stranded β barrel, the trimeric integral outer-membrane porin OmpG from *E. coli*, PDB ID 2f1c.

https://doi.org/10.1371/journal.pone.0180905.g015

see Fig 17 and S4 Appendix. The network of backbone-backbone H-bonding in PDZ domains also appears to fit into this matrix, see Fig 14 and S3 Appendix.

(vi) Stability of tertiary structure and metamorphic equilibria: Refolding upon reorganization of the folding basin *FB*. The preceding discussion implies in what circumstances one amino acid sequence can support multiple folds. The 'C_5 strand', 'C_{7eq} strand' and 'turn' segments that are incorporated into the α-helices to organize and stabilize globular tertiary structure or to ensure function-required mobility, may form bona fide elements of a stable β structure. On the other hand, the 'helix' segments may form the 'C_{7eq} strands' in a non-polar environment or the 'C_5 strands' in a highly polarizing environment. All such relatively short segments would be identified as the *chameleon* sequences [142].

A much longer polypeptide chain can adopt more than one well-structured fold when a molecule comprises 'malleable' backbone segments which (1) have ambiguous conformational propensities (lack strong FP_i bias), (2) are amenable to the polarizing/depolarizing effects of the molecular embedding and the medium, and (3) are connected by a 'hinge' that controls their mutual orientation. A change in the conformation of the 'malleable' segment may then occur as a result of reorganization of the protein's folding basin *FB* brought about by a change in the conformation of the 'hinge'. When these two conformational transitions are coupled, the protein may exist in the metamorphic equilibrium [37]. The concept of the 'hinge'-controlled reorganization of the folding basin *FB* and the concomitant metamorphic equilibration is illustrated in Fig 18.

In the first three cases [143–145], Fig 18A–18C, the common feature is the C-terminal segment which has 'helical' $\Delta FP_{i-1 \to i+1}$ profile but *depressed* $\overline{FP_i}$ (≤ 0) and may or may not be in close contact with the N-terminal subdomain depending on the conformation of the preceding
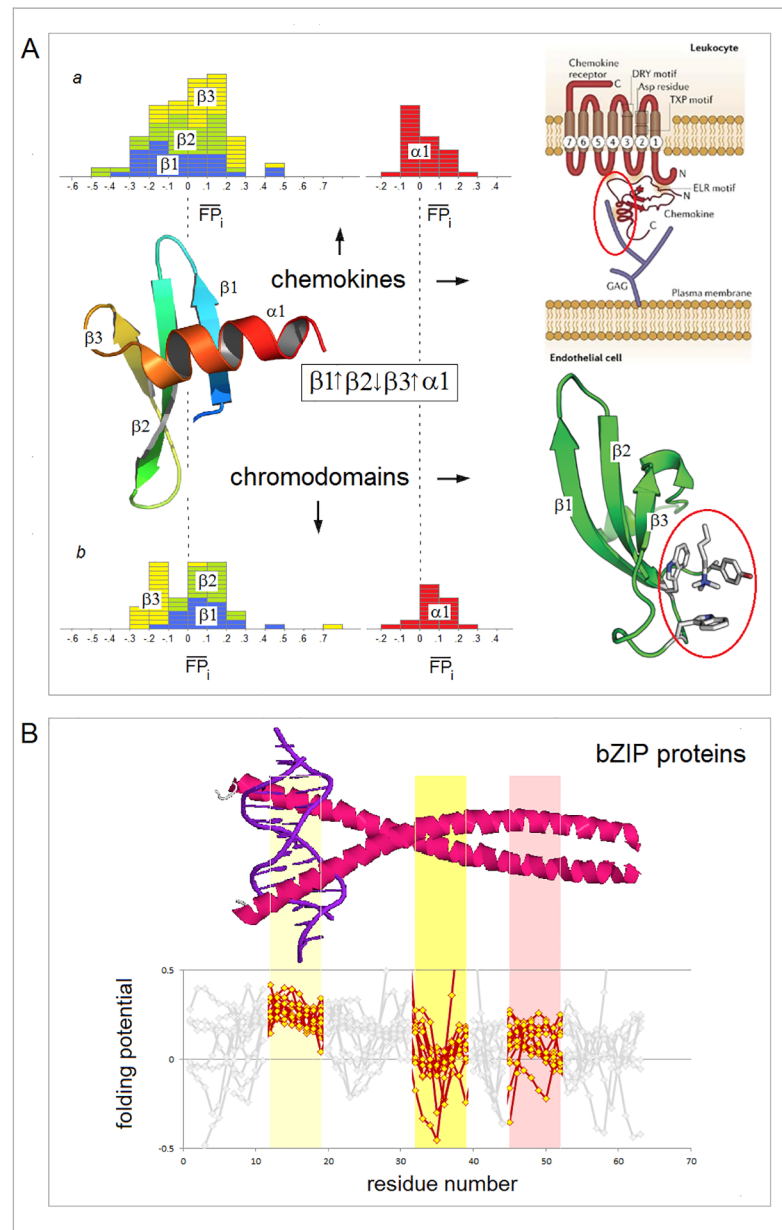
Fig 16. Electronic configuration of the polypeptide backbone versus complexity of tertiary structure and order of oligomerization. (A) The histograms compare the distribution of the α-helix and β-sheet $\overline{FP_i}$ in the chemokine and CHROMO domains [140,141]. The βββα folds of these two domains are very similar but the roles of the N-terminal and C-terminal segments are reversed: the chemokines use the C-terminal helix to bind the substrates while the CHROMO domains use the N-terminal β-hairpin. The expected difference in the $\overline{FP_i}$ distribution is indeed found (β1 cyan, β2 light cyan, β3 yellow, α1 red): (a) In chemokines, the binding function involves the C-terminal β3 and α1, and indeed the N-terminal β1 and β2 are the archetypal 'C$_5$ strands' ($\overline{FP_i}$ largely <0) and anchor the structure while the C-terminal α1 will only be stable when it is partially buried ($\overline{FP_i}$ tends to be <0, PDB IDs: 1b3a, 1cm9, 1dok, 1eih, 1el0, 1eot, 1esr, 1f2l, 1f9p, 1g2t, 1g91, 1ha6, 1il8, 1j9o, 1m8a, 1mgs, 1ncv, 1nr4, 1o80, 1qg7, 1qnk, 1rhp, 2hcc, 2kol, 2kum, 2l4n, 2q8r, 2ra4, 3n52, 3ona, 3tn2, 4hcs, 4hsv); (b) In the CHROMO domains (binding involves the N-terminal β1 and β2), the C-terminal α1 is the archetypal 'helix' (0<$\overline{FP_i}$<0.2) and anchors the structure while the N-terminal β1 and β2 strands (also 0<$\overline{FP_i}$<0.2) will be stabilized when they are either partially buried and therefore less polarized (turning C$_{7eq}$), or attached to a cross-β structure and therefore more polarized (turning C$_5$) (PDB IDs: 1ap0, 1g6z, 1kna, 1pdq, 2b2v, 2d9u, 2dnt, 2dnv, 2dy8, 2ee1, 2epb, 2fgg, 2k1b, 2rsn, 2rso, 3fdt, 3gv6, 3i91, 3mts, 3r93, 3tzd). **(B)** The FP_i plots for bZIP proteins (PDB IDs: 1ci6, 1dgc, 1dh3, 1gd2, 1jnm, 1nwq, 1t2k, 1s9k, 2c9l, 2e43).

The multiple alignment suggests that $\overline{FP_i}$ and function of each distinct segment of bZIP helix are related. Based on the assignments in Fig 7(c), the basic N-terminal segment has the 'C_{7eq} strand' $\overline{FP_i}$, the mid-region segment has the 'C_5 strand' $\overline{FP_i}$, and the C-terminal segment has the 'helix' $\overline{FP_i}$. Thus, the folding potential FP_i of each segment appears optimized to ensure stability of the helical fold in three different environments created in the complex of bZIP dimer and DNA duplex.

'hinge'. Consequently, this segment folds into the α-helix in a low-permittivity environment i.e. when the C- and N-terminal subdomains are brought into close contact and the folding basin FB increases in size, cf. section **b**.(iv); otherwise, it may be incorporated into β structure or remain unstructured when exposed to solvent. In the fourth case, Fig 18D, the 56-residue protein switches folds back and forth between β grasp (ββαββ) and helix-bundle (ααα), each time as a result of replacement of a single residue in one of the two 'hinges' in the molecule [146,147]. In the β grasp fold, Fig 18D(a), the N-terminal subdomain of this protein forms the 'C_{7eq} strands' and β1-β2 hairpin, even though it lacks the 'turn' segment and has distinctly 'helical' $\Delta FP_{i-1\to i+1}$ profile and optimally 'helical' $\overline{FP_i}$. To stabilize this fold: (1) the β1-β2 hairpin is placed in a low-permittivity environment by the contact with the mid-region α-helix buried on its surface, which requires that the connecting 'hinge' (~residues 20–25) is sufficiently long and stable; and (2) the β1-β2 hairpin is incorporated into extended β structure via parallel $C_{7eq}\uparrow C_5\uparrow$ alignment with the C-terminal β3-β4 hairpin. Thus, the A20L mutation, (a)→(b) in Fig 18D, apparently shortens the 'hinge' so that the helical propensity of the N-terminal segment prevails and prompts the transition to the three-helix-bundle fold; the compact all-α fold is stabilized by the incorporation of the 'C_5 strand' segments into the helices α2 and α3. The I25T mutation, (b)→(c) in Fig 18D, stabilizes the required bend (increases the steep *change* in FP_i in the segment 20–25) and restores the β grasp fold. This fold, however, cannot be stable without the stable C-terminal $C_5\uparrow C_5\downarrow$ hairpin, and so the Y45L mutation that abolishes the C_5 propensity of β3 strand also prompts the conversion into the helix bundle, see (c)→(d) in Fig 18D.

(vii) Stability of tertiary structure and metamorphic equilibria: Refolding of β structure that involves reorganization of the folding basin *FB* and *FP*-directed molecular recognition. The last example in Fig 18 is the refolding of β structure of the spindle checkpoint protein Mad2 [148]. Mad2 comprises the 'fixed' {β3↑β1↓β2↑}-sheet and the 'malleable' C-terminal β-hairpin {β4↑β5↓} attached by an unstructured loop. The interconversion between two distinct folded conformations of Mad2 involves reconfiguration of the 7-residue segment (S185-H191) which has highly negative $\overline{FP_i}$ ($\ll 0$). Based on the assignments in Fig 7, these 7 residues may adopt the 'C_5 strand' configuration in a low-permittivity environment i.e. when they are buried, and the '$FP_i \ll 0$ turn' configuration when they are exposed to solvent. When such a configuration switch occurs, the C-terminal hairpin shifts its position along the chain, changing loop's length and 'strand' propensities: it forms either (1) the 'short loop'-$C_5\uparrow C_5\downarrow$ hairpin, Fig 18E(a), or (2) the 'long loop'-$C_{7eq}\uparrow C_{7eq}\downarrow$ hairpin, Fig 18E(b). The shift reorganizes the folding basin *FB* because it determines how the {β3-β1-β2} and {β4-β5} fragments are attached to each other. In the first case, the arrangement is {β4-β5}-{β3-β1-β2}: the S185-H191 residues are buried as a part of the β5 'C_5 strand' which is parallel to the proximal β3 (the preferred alignment $C_{7eq}\downarrow C_5\downarrow$ for β3-β5 cf. Figs 2 and 5). In the second case, the arrangement is {β3-β1-β2}-{β5-β4}: the S185-H191 residues are exposed to solvent as the 'turn' of the β4-β5 hairpin, while the β5 'C_{7eq} strand' is antiparallel to the distal β2 (the preferred alignment $C_{7eq}\uparrow C_{7eq}\downarrow$ for β2-β5). Thus, the reorganization of Mad2 β-sheet appears to be a case of the *FP*-directed molecular recognition. The difference in the β-sheet organization is also consistent with the difference in the loop's reach (length).
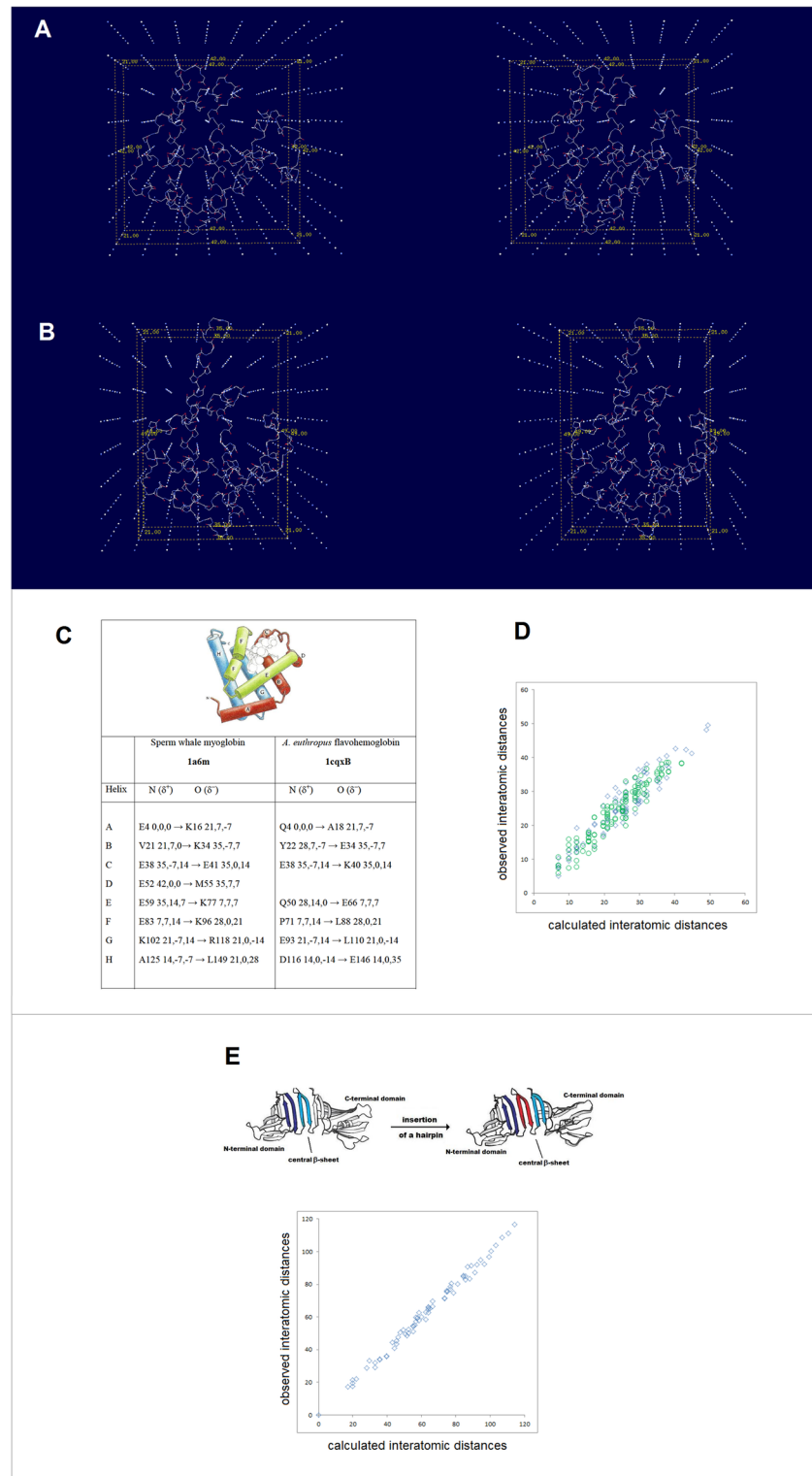
Fig 17. Folding template and tertiary structure of soluble globular proteins. (A) The interaction of sperm whale apomyoglobin, PDB ID 1a6m, with the transient cubic lattice of the ionic atmosphere in the physiological 1:1 electrolyte solution: mutual polarization of the polypeptide backbone and the Ghosh-Debye-Hückel matrix stabilizes the protein/electrolyte system when the ends of the helices A-H of the globin fold replace the corresponding salt ions in the vertices of the matrix with the 7 Å lattice constant. The 'template

box' has dimensions 42×42×21 (Å). **(B)** In the case of *Alcaligenes euthropus* flavohemoglobin, PDB ID 1cqxB, the 'template box' has dimensions are 49×35×21 (Å) and it is polar along each Cartesian axis and so it is chiral. **(C)** Manual search for the best fit of the termini of the A-H helices of both globins into the Ghosh-Debye-Hückel matrix yields the coordinates shown in this panel (partial charges δ^+ and δ^- of the helix termini are represented by the N and O atoms of the residues assigned as the terminal helix residues by the Swiss-PDBViewer with one exception of the N terminus of the helix B in 1a6m where PDBViewer assignment is erroneous; the Cartesian coordinate system is left-handed). **(D)** The calculated and the observed matrices of the interatomic distances (Å) between the helix termini in: ○ sperm whale apomyoglobin, PDB ID 1a6m, and ◊ *Alcaligenes euthropus* flavohemoglobin, PDB ID 1cqxB (based on the coordinates listed in the panel C). **(E)** The calculated and the observed matrices of the interatomic distances (Å) between the N^ξ atoms of the 'reporter' lysines in the congeners of *Borrelia* spirochete antigen OspA (PDB IDs 2af5, 2fkg, 2fkjA, 2fkjB, 2fkjC, 2g8c, 2hkd, 2i5v, 2i5z, 2ol6, 2ol7A, 2ol7B, 2ol8, 2oy1, 2oy5, 2oy7, 2oy8, 2oyb, 2pi3, 3ckaA, 3ckaB, 3ckf, 3ckg, 3ec5, 3eexA, 3eexB). See S4 Appendix for the definition of the plotted data.

(viii) Coupled folding and binding: Formation of α structure upon reorganization of the folding basin *FB*. Molecular recognition features (MoRFs) [38,149], just like the 'malleable' segments of metamorphic proteins, can change conformation as a result of mere change in environment's capacity to polarize the polypeptide backbone. Transfer of a disordered protein segment from water into a complex with another protein is one way to achieve such as change, see Fig 19. The FP_i vs. $\Delta FP_{i-1\rightarrow i+1}$ plots show distinctly helical $\Delta FP_{i-1\rightarrow i+1}$ profiles for the activation domain of CITED2 and the α-MoRF of the measles virus N protein but neither can fold in solution since $\overline{FP_i}$ is too low in the first case and too high in the second case, cf. Fig 19A(a) and 19B(a) [150–152]. To form a well-folded helix, the CITED2 domain needs to be buried in a protein ligand i.e. placed in a low-permittivity environment, and this is indeed found in the complex with the TAZ1 domain of the transcriptional coactivator CREB-binding protein (CBP) (Chart 11A(b), PDB ID 1r8u) [150]. In contrast, the viral MoRF, which forms a complex with the nucleocapsid-binding domain of the measles virus P protein (PDB ID 1t6o), remains bound in the polarizing environment of the protein surface/ionic atmosphere shell, Fig 19B(b) [151,152]. The corresponding difference in the organization of tertiary structure of the two all-α ligands is quite apparent in the $FP_i/\Delta FP_{i-1\rightarrow i+1}$ plots, see Fig 19A(c) and 19B(c). However, both complexes seem optimally set up for the interaction with the transient lattice of ionic atmosphere at the physiological conditions, see Fig 19A(d) and 19B(d).

(ix) Coupled folding and binding: Formation of β structure via *FP*-directed molecular recognition. Coupled folding and binding may also involve formation of β structure. Autonomous folding of a stable β sheet requires precise location and matching of distinct 'strand' and 'turn' propensities, see Figs 2 and 5. Thus, a polypeptide chain which comprises mismatched or lacking strong FP_i bias 'strand' and 'turn' segments will remain disordered unless the *anticipated* 'strand' configurations are stabilized in an extended intramolecular or intermolecular β structure. The synergistic folding of split inteins reveals a mechanism of molecular recognition that is likely to underlie such processes, in this case formation of the two-stranded antiparallel β-sheet $\beta_N\uparrow\beta_C\downarrow$ that anchors the horseshoe-like fold [153]. The β_N and β_C segments are 30-35 residue-long and remain disordered in solution but form upon binding an extensive left-handed coiled-coil which is a variation of the common fold of two-stranded antiparallel sheets expected to display 'extraordinary flexibility' [154]. In inteins, the coiled-coil has a non-uniform twist so that each strand comprises four distinct segments, $\beta_N a'-\beta_N a-\beta_N b-\beta_N b'$ and $\beta_C a'-\beta_C a-\beta_C b-\beta_C b'$, and has, in addition, a left-handed superhelical twist, Fig 20A. This architecture appears consistent with the intrinsic right-handed twist of the extended strands, distribution of secondary-structure propensities and insertion of turn sequences, and length limitations on the cooperativity of backbone-backbone H-bonding [155]. According to our model, see Figs 2 and 5, molecular recognition in the assembly of the $\beta_N\uparrow\beta_C\downarrow$ sheet should be
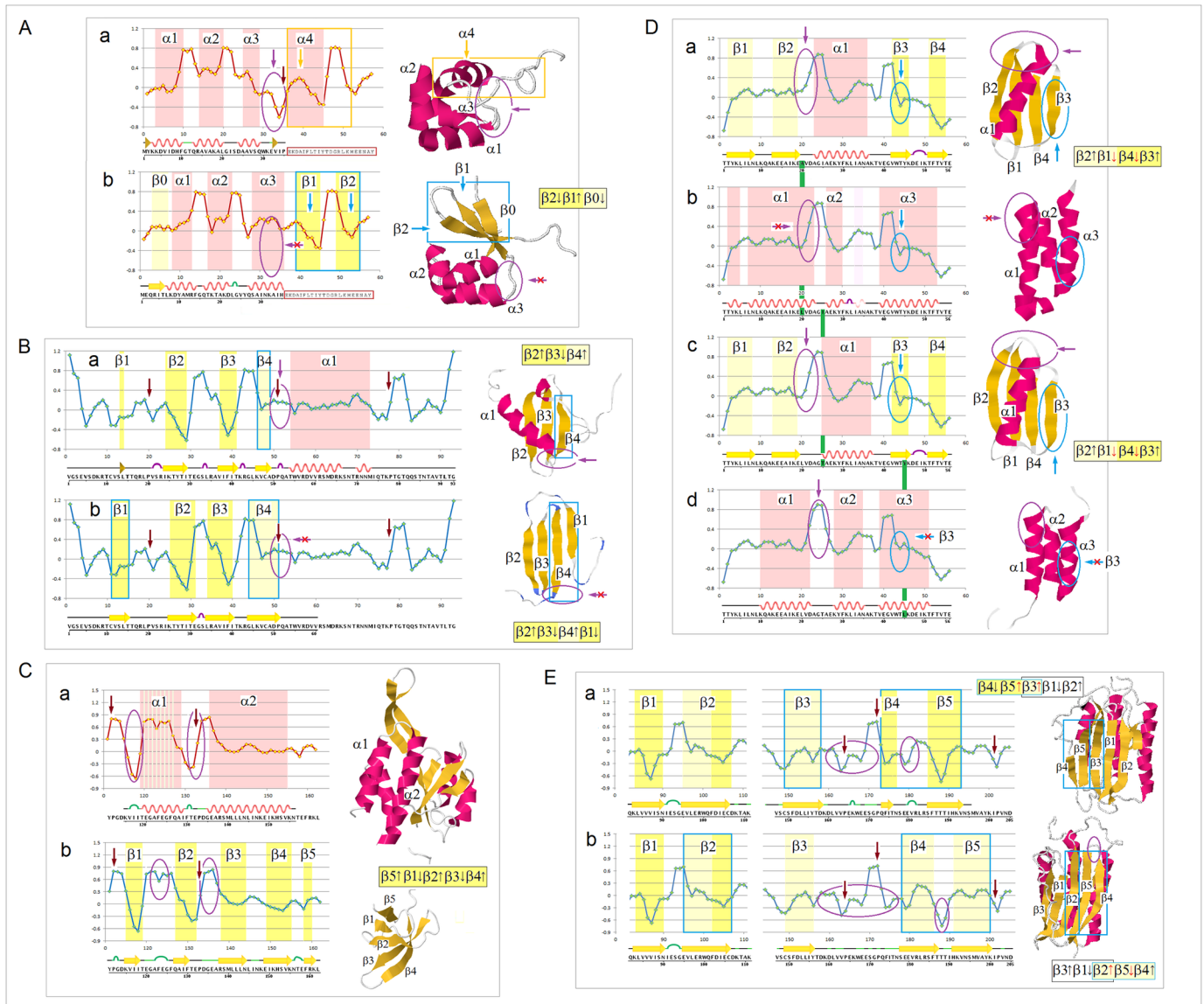
Fig 18. Electronic configuration of the polypeptide backbone and metamorphic equilibria. (A) Chameleon sequence SASF2 attached to N-terminal segments of λ Cro and p22 Cro repressors. When attached to the 35-residue N-terminal segments of p22 Cro (PDB ID 1rzs) or λ Cro (PDB ID 5cro) repressors, the sequence SASF2 mimics the conformations of the native C-terminal segments of these repressors, folding into an α-helix in the first case and into a β-hairpin in the second case [143]: (a) The FP_i plot for the p22 Cro/SASF2 conjugate. The chameleon sequence SASF2 is preceded by a flexible 'hinge' formed by the short C_5 strand (W30-E32, $\overline{FP_i} = -0.1080$) and the PP_{II} helix (V33-P35, $\overline{FP_i} = -0.4542$), which extends back towards the helices α1, α2 and α3 (PDB ID 1rzs) so that the main chain makes U-turn placing SASF2 in the helix bundle; (b) The FP_i plot for the λ Cro/ SASF2 conjugate. The segment preceding SASF2 is the long helix α3 (S28-H35, $\overline{FP_i} = 0.1787$) (PDB ID 5cro) and the backbone cannot make U-turn. **(B) Human lymphotactin.** Under physiological conditions, human lymphotactin exists in the equilibrium between the canonical chemokine monomer PDB ID 1j8i and the homodimer of the Greek-key β-sheet PDB ID 2jp1 [144,145]: (a) The FP_i plot for the monomeric chemokine-like conformer. The misalignment of the antiparallel C_5 and C_{7eq} strands (β3 and β4) creates a flexible 'hinge' at the end of β4 so that the low-$\overline{FP_i}$ helix α1 can bury itself on the concave surface formed by the β structure and the unstructured, 'S-S attached' N-terminal segment; (b) The FP_i plot for the dimeric Greek-key conformer of lymphotactin. The stabilization of α1 is not possible when the extension of the three-stranded β2↓β3↑β4↓ sheet into the four-stranded one β2↓β5↓β4↓β1↑ obliterates the 'hinge' (β4 is lengthened due to additional alignment of β1) and flattens the lymphotactin surface (the N-terminal segment folds into β1). **(C) E. coli transcription factor RfaH.** The N-terminal (NTD) and C-terminal (CTD) domains of RfaH are connected by a flexible linker which allows either for tight interaction or complete separation of the two domains. The arrangement determines the fold of CTD: (a) The FP_i plot for the CTD partially buried in the NTD, PDB ID 2oug; (b) The FP_i plot for the autonomous CTD isolated in the aqueous environment, PDB ID 2lcl. **(D) G$_B$98/G$_A$98 variants of the streptococcal albumin- and immunoglobulin-binding proteins.** The FP_i plots for the 56-residue GB1 congeners and the alternation of folds produced in each case by the replacement of a single residue (mutation sites are marked by the green vertical bars) [146,147]: (a) The FP_i plot for the variant PDB ID 2lhe (G$_B$98-T25I/L20A); (b) The FP_i plot for the variant PDB ID 2lhg (G$_B$98-T25I); (c) The FP_i plot for the variant PDB ID 2lhd (G$_B$98); (d) The FP_i plot for the variant PDB ID 2lhc (G$_A$98≡G$_B$98-Y45L). **(E) Human mitotic spindle protein Mad2**

[148]: (a) The FP_i plot for the five-stranded β sheet of Mad2 with the C-terminal hairpin in the $C_5{\downarrow}C_5{\uparrow}$ configuration, PDB ID 1duj; (b) The FP_i plot for the five-stranded β sheet of Mad2 with the C-terminal hairpin in the $C_{7eq}{\downarrow}C_{7eq}{\uparrow}$ configuration, PDB ID 1s2h.

https://doi.org/10.1371/journal.pone.0180905.g018

aided by the 'symmetry' of backbone polarization of β_N and β_C, in particular by 'matching' of the middle-segment pairs $\beta_N a$-$\beta_N b$ and $\beta_C a$-$\beta_C b$. The FP_i plots for the 15 inteins of known three-dimensional structure, see Fig 20B–20D, do show such a trend: the complete 'symmetry' of backbone polarization of the $\beta_N{\uparrow}\beta_C{\downarrow}$ middle segments is found in 9 structures, partial 'symmetry' in 2 structures (split inteins), and none in 4 structures, which yields the total of 20 FP_i-'matching' and 10 FP_i-'mismatched' middle-segment pairs. In principle, complementarity of the electrostatic charges carried by β_N and β_C could also aid folding of this coiled-coil as the folding-template effect, and it was indeed invoked to explain the kinetics of split-intein folding [156]. According to the present data, such assistance does appear to operate in some cases but it is not required: the β_N and β_C strands carry complementary charges in 11 cases, in one case both strands carry negative charge, and one or both strands are neutral in 3 cases. However, in the split inteins PDB IDs 1zd7 and 2keq, complementarity of *multiple* charges is found in the $\beta_N b$-$\beta_C a$ segment pairs that lack the 'symmetry' of backbone polarization while in the symmetry-wise 'matching' $\beta_N a$-$\beta_C b$ segment pairs both segments are either neutral (1zd7) or both carry negative charges (2keq), cf. Fig 20C. This trend is reproduced in the total sample: among the 20 'matching' $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$ pairs, both segments carry *the same charge* in 4 cases, in 11 cases one of the two segments is neutral, and the charges are complementary in 5 cases (with just a single charge on one of the segments), Fig 20B and 20C. On the other hand, among the 10 middle-segment pairs that lack the 'symmetry' of backbone polarization, the charges are complementary in 8 cases (with *multiple* charges on both segments in 4 cases), and only in 2 cases one of the two segments is neutral, Fig 20C and 20D. These findings are consistent with the notion that the side chain-side chain interactions may facilitate the initial stages of intein folding via the 'electrostatic capture' [156], or stabilize the 3D structure once the backbone fold is established, but neither effect is necessary to assembly the two-stranded antiparallel β-sheet of inteins in the correct register. The discrete architecture of the $\beta_N{\uparrow}\beta_C{\downarrow}$ sheet seems to depend on the complex pattern of backbone polarization, and the 'symmetry' of polarization of the $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$ segments appears to be an important feature of this pattern. For the detailed account of the discussed data see S5 Appendix.

(x) pH-Driven equilibria: Transformation of secondary structure. Conformational propensity of a backbone segment may be altered by the change in pH. For instance, amylin adopts β-hairpin fold in a neutral or basic solution where the C-terminal segment has the $\Delta FP_{i\text{-}1\rightarrow i+1}$ profile 'C_5 strand'-'turn'-'C_5 strand', but it becomes helical in acidic solution [157] where, due to the protonation of His-18, the entire N-terminal segment acquires 'helical' $\Delta FP_{i\text{-}1\rightarrow i+1}$ profile (T9-N22: $\overline{FP}_i = 0.1746$), Fig 21A. On the other hand, the B-loop of influenza hemagglutinin has a 'helical' $\Delta FP_{i\text{-}1\rightarrow i+1}$ profile but very low $\overline{FP}_i$ and is therefore unstructured at pH>7, cf. Fig 21B. At low pH, however, B-loop residues do form a long helix [158] and indeed the FP_i vs. $\Delta FP_{i\text{-}1\rightarrow i+1}$ plot at pH<5 reveals a considerable shift of its 'helical' cluster to $FP_i > 0$.

(xi) pH-Driven equilibria: Loss of tertiary structure and reorganization of the folding basin *FB*. In the case of the α-pore forming diphteria toxin, the crucial unfolding of several α-helices of its translocation domain T is triggered by the reduced pH in the endosomal lumen (pH 4.5–5). By attenuating polarization of the polypeptide backbone, low pH may increase mobility of the α-helices of the diphteria toxin and 'neutralize' the 'C_5 strand' segments which ought to be incorporated into these α-helices to stabilize the compact tertiary structure of the
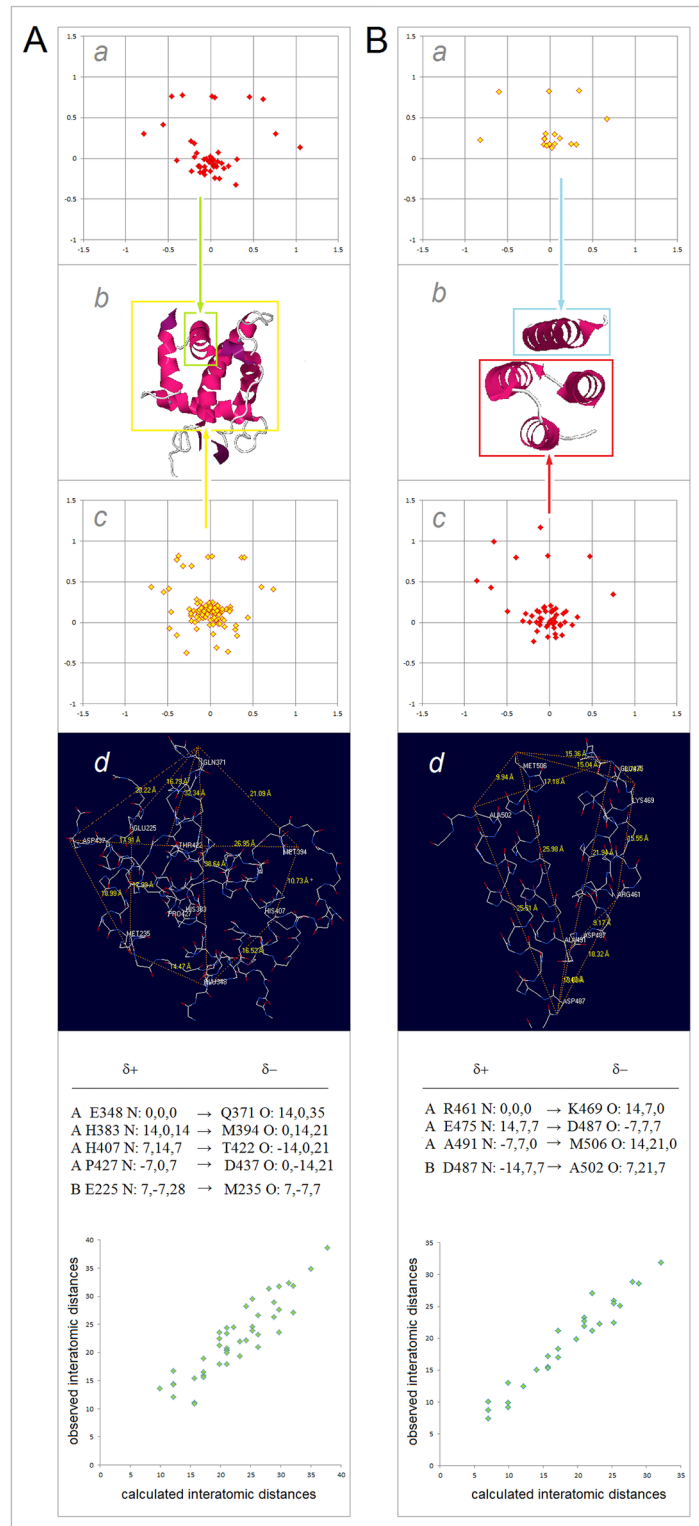
Fig 19. Electronic configuration of the polypeptide backbone and molecular recognition in formation of α structure: Folding of α-MoRFs [149–152]. (A) The complex between the TAZ1 domain of the transcriptional coactivator CREB-binding protein (CBP, the 4-helix-bundle IDP receptor) with the activation domain of CITED2 (IDP), PDB ID 1r8u [150]: (a) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plot for the activation domain of CITED2. The domain is unstructured in the free form in aqueous solution as expected given its $FP_i/\Delta FP_{i-1 \to i+1}$

pattern: FP_i<0; (b) The N-terminal segment of the activation domain of CITED2 is helical in the complex with TAZ1 where it is deeply buried (cf. the helix within the green outline) while the C-terminal segment, exposed to solvent, remains unstructured; (c) The FP_i vs. $\Delta FP_{i-1\to i+1}$ plot for the CBP TAZ1 domain shows pronounced archetypal 'helix' propensity, cf. Fig 5. Note that the α-helices of this domain do not incorporate 'C$_5$ strand' segments which would be expected to stabilize the bundle structure—in fact, the assembly is stabilized by the chelation of three Zn^{2+} ions; (d) The projected fit of the putative key surface charges δ^+ and δ^- (the termini of the five helices shown in the Swiss-PDBViewer projection) into the Ghosh-Debye-Hückel matrix, and the correlation between the predicted by the folding-template model and the observed interatomic distances (Å) between the charges δ^+ and δ^-. **(B)** The complex between the α-MoRF of the measles virus N protein and nucleocapsid-binding domain of the measles virus P protein, PDB ID 1t6o [151,152]: (a) The FP_i vs. $\Delta FP_{i-1\to i+1}$ plot for the α-MoRF. The FP_i profile is helical (approximately constant FP_i) but $\overline{FP_i}$~0.3 is too high for the aqueous environment and the domain is disordered in solution; (b) Given its $FP_i/\Delta FP_{i-1\to i+1}$ pattern, the measles virus α-MoRF is expected to fold into a helix not when it is buried but when it remains in the solvent shell of the binding protein, benefiting from the synergistic stabilization of the ionic matrix, and this is indeed found in the complex, cf. the helix in the blue outline; (c) The FP_i vs. $\Delta FP_{i-1\to i+1}$ plot for the nucleocapsid-binding domain shows pronounced 'helix' propensity characteristic of helix bundles, cf. Fig 15A; (d) The projected fit of the putative key surface charges δ^+ and δ^- (the termini of the four helices shown in the Swiss-PDBViewer projection) into the Ghosh-Debye-Hückel matrix, and the correlation between the predicted by the folding-template model and the observed interatomic distances (Å) between the charges δ^+ and δ^-.

T domain bundle, cf. sections **a**.(iii) and **b**.(iv). The FP_i vs. $\Delta FP_{i-1\to i+1}$ plots in Fig 21C suggest that this is indeed the underlying effect. In particular, the folding potential FP_i of the two helices that completely unfold under the acidic conditions, TH3 and TH4 [159], shifts from $FP_i \leq 0$ (the α-helices which are stable in the interior of a compact structure) to $FP_i >0.3$ (the α-helices which are unstable except in a highly polarizing environment), see Fig 21C(d) and 21C(e).

It follows that the shift in backbone density distribution produced by protonation of D^-, E^- and H may destabilize compact tertiary structure and trigger the transition into an 'acid molten globule' which can in turn initiate fibrillization of a globular protein. The acid-induced unfolding and fibrillization of transthyretin illustrates this point. Based on the FP_i plots, several segments of the polypeptide backbone of transthyretin considerably change conformational and H-bonding propensity at pH 4, see Fig 21D(a): (1) the FP_i profile in the βB region becomes 'helical' due to the protonation of His-31, cf. similar 'strand'→'helix' transition in amylin; (2) the pronounced 'C$_5$ strand' propensity of the βC, βE and βF strands is lost; (3) the high 'turn' propensity of the DE loop (D^* segment) is lost (the C$_5$-like, negative $\overline{FP_i}$ of D^* becomes positive and C7$_{eq}$-like); (4) the high 'turn' propensity of the EF loop (E^* segment) is lost [160,161]. These changes destabilize the outer β sheet of transthyretin and lead to step-by-step unravelling of the β sandwich fold beginning with 'unmasking' of the βA strand. Subsequent 'unmasking' of the βB strand, or alternatively the extrusion of the βB strand and insertion of the βD* strand in its place, generate three other partially folded monomers as shown in Fig 21D(b). The seemingly contradictory results of the solid state NMR [162–164], spin-labelling [165,166], immunoreactivity [167], and H/D exchange (HXMS) [168,169] investigations can be reconciled by assuming that all four putative aggregation intermediates are competent to assemble into annular octamers [170–172], protofibrils and fibrils, see Fig 21D(c), and the actual course of fibrillization depends on the incubation conditions.

c. The PMO theory of misfolding and aggregation of pleiomorphic proteins associated with neurodegenerative diseases

Misfolding of several proteins is believed to play a key role in genesis of neurodegenerative diseases including Alzheimer's amyloid β (Aβ) and tau proteins, Parkinson's α-synuclein and TSEs' mammalian prions. These proteins are highly pleiomorphic and readily undergo a series of conformational transitions along the aggregation pathways [173]. Based on the preceding considerations and findings, and taking into account extensive experimental evidence [173–

Fig 20. Electronic configuration of the polypeptide backbone and molecular recognition in formation of β structure: Folding of inteins [153,156]. The FP_i assignment of secondary-structure propensity vs. distribution of side chain charges in the two-stranded antiparallel β-sheet $\beta_N\uparrow\beta_C\downarrow$ of 15 inteins and intein-like domains of the known 3D structure. **(A)** The architecture of the horseshoe-like fold of inteins: (a) the antiparallel assembly of the 30–35 residue-long strands β_N and β_C into the $\beta_N\uparrow\beta_C\downarrow$ sheet constitutes the major

feature of the HINT fold. Each strand comprises four segments, $\beta_N a'$-$\beta_N a$-$\beta_N b$-$\beta_N b'$ and $\beta_C a'$-$\beta_C a$-$\beta_C b$-$\beta_C b'$ (cf. the labels color-coded blue and red, respectively), which are marked by the changes in the direction of the main chain (produced by the inherent right-handed twist of the polypeptide backbone and the insertion of turns); (b) the putative assembly of the full-length β_N and β_C strands into the $\beta_N\uparrow\beta_C\downarrow$ sheet. The interactions of the terminal-segment pairs, $\beta_N a'$-$\beta_C b'$ and $\beta_N b'$-$\beta_C a'$, may contribute to the stabilization of the correct register in the initial stages of folding; (c) in the native state, the $\beta_N\uparrow\beta_C\downarrow$ sheet comprises the middle-segment pairs $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$ since the C-terminal segments of both strands, $\beta_N b'$ and $\beta_C b'$, are folded towards the protein interior. **(B)** The FP_i assignment of secondary-structure propensity vs. distribution of the side chain charges in the case of the full 'symmetry' of backbone polarization of the middle-segment pairs $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$, shown within the green rectangles in the schematic representation of the $\beta_N\uparrow\beta_C\downarrow$ sheet. The PDB ID-labeled diagrams show FP_i plots, DSSP assignments, and the sequences of the 30–35 residue-long β_N and β_C strands with D, E marked by red squares and K, R marked by blue squares. The four segments of each β_N and β_C strand are defined based on the DSSP/Swiss-Prot assignments and the inspection of the 3D structures, and are shaded yellow or light-yellow in the FP_i plots depending on the secondary-structure preferences; the latter are assigned based on Figs 5 and 7. **(C)** The FP_i assignment of secondary-structure propensity vs. distribution of side chain charges in the case of the partial 'symmetry' of backbone polarization of the middle-segment pairs $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$. The 'matching' $\beta_N a$-$\beta_C b$ pairs are shown within the green rectangles and the 'mismatched' $\beta_N b$-$\beta_C a$ pair within the red rectangles in the schematic representation of the $\beta_N\uparrow\beta_C\downarrow$ sheet. **(D)** The FP_i assignment of secondary-structure propensity vs. distribution of side chain charges in the absence of the 'symmetry' of backbone polarization of the middle-segment pairs $\beta_N a$-$\beta_C b$ and $\beta_N b$-$\beta_C a$, shown within the red rectangles in the schematic representation of the $\beta_N\uparrow\beta_C\downarrow$ sheet. For a detailed report on the discussed data see S5 Appendix.

https://doi.org/10.1371/journal.pone.0180905.g020

370], we develop here a model of pleiomorphism of these proteins focusing initially on the best-documented case, the polymerization of Aβ.

   (i) Antiparallel coiled-coil as a model of Aβ dimer structure.   The Aβ fold strongly depends on medium and molecular embedding, and the FP_i plots in Fig 22A correctly anticipate this dependence: secondary structure propensities assigned according to Fig 7 are consistent with the available data [174–194]. The 'helix' and 'strand' segments of Aβ lack the FP_i bias needed to stabilize the α or β structure in water, Fig 22A(c), and indeed the monomeric Aβ exists in water as a collapsed coil devoid of secondary structure. However, the anticipated 'strand' configurations of Aβ may be stabilized in an extended intermolecular β sheet i.e. as a result of oligomerization. Since the experimental evidence points to the importance of the simplest oligomer, the Aβ homodimer, in aetiology of neurodegeneration [195–199], we assume that this dimer is the basic building block of polymerization. The SAXS data for the stoichiometric $A\beta^{1-42}$ complexes with 8-hydroxyquinolines suggest that such complexes involve specific extended dimers of $A\beta^{1-42}$ which are at least partially folded [200], Fig 22B(a). The technique does not allow to establish the mode of alignment; we propose below how such extended dimers are aligned by invoking the FP-directed molecular recognition, cf. sections **a**.(i) and **b**.(vii) and (ix), and the solid-state NMR data for the off-pathway Aβ oligomers 150±30 kDa [201,202].

   Fig 22B(b) shows the 'strand' configurations in the homodimer that favor fibrillization [203,204]. The sequence $\beta 0(C_{7eq})$-$\beta 1(C_5)$-$\beta 2(C_5^*)$-$\beta 3(C_5)$ is crucial to the FP-directed molecular recognition in dimerization and subsequent pleiomorphic behavior of Aβ. Some of the observations consistent with this proposition are the following: (1) the FAD English H6R, Taiwanese D7H and Tottori D7N mutations which promote aggregation [205,206] would stabilize the $C7_{eq}$ conformation of β0 while the A2T and A2V mutations which also influence aggregation [207–209] would stabilize the C_5 fold; (2) the CHC mutations that stabilize the C_5 conformation of β1 (e.g. F19I/F20I shown here in the FP_i plot) appear to increase the rate of primary nucleation [210–213], although the evidence is not unequivocal [214], while the A21G mutant (Flemish FAD) does not undergo standard fibrillization [215]; (3) conservative hydrophobic mutations that preserve low permittivity environment required to stabilize the $\beta 2(C_5^*)$ H-bonding, cf. section **a**.(iii), do not seem to affect the initial oligomerization [216]; (4) VPV
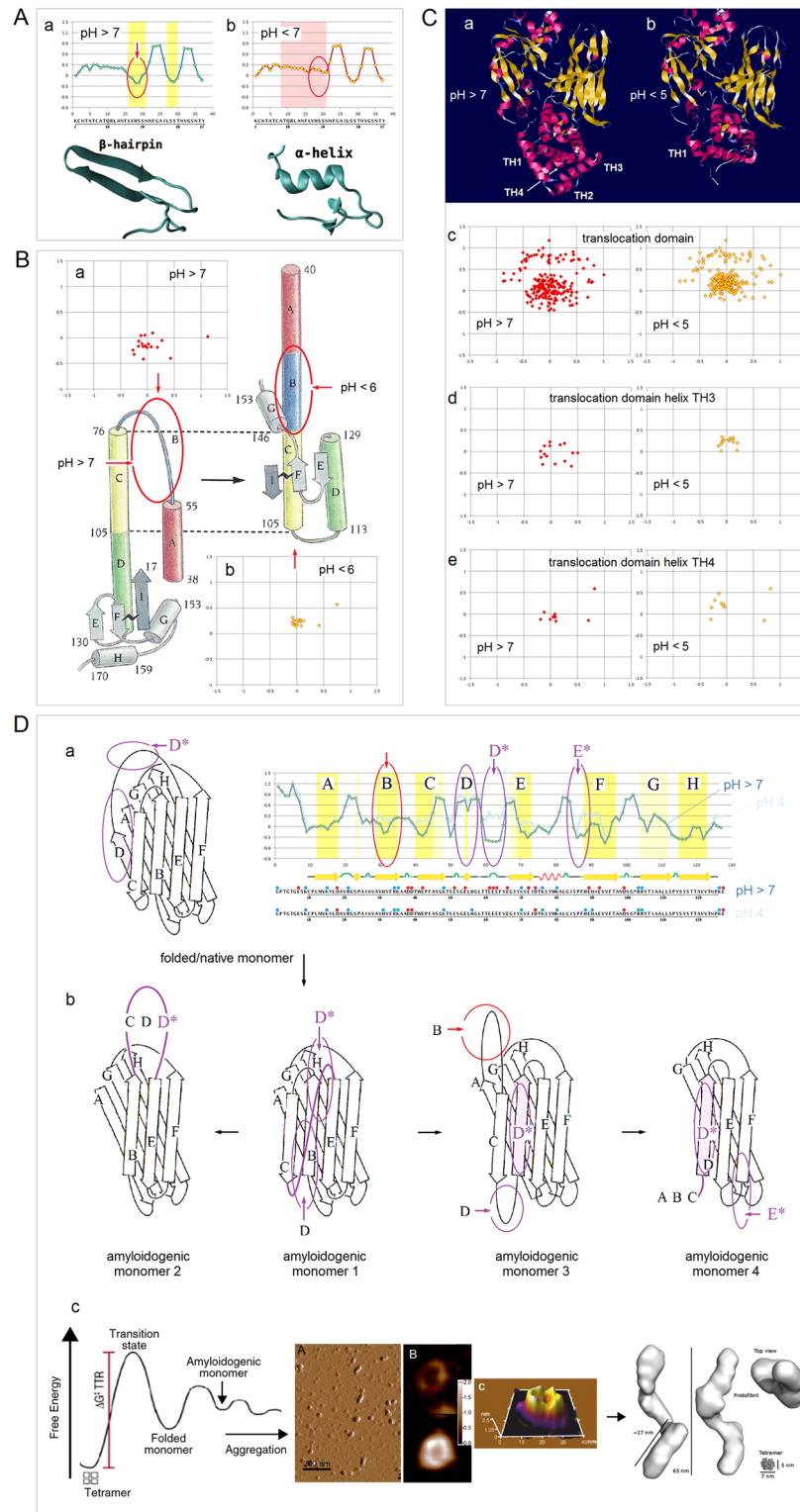
**Fig 21. Effect of pH on conformational and H-bonding propensity of the polypeptide backbone. (A)
Conformational equilibria of islet amyloid polypeptide (amylin)** [157]: (a) The FP_i plot of amylin at high
pH and the putative β-sheet conformation of the peptide. Assuming $\sigma^{His} \equiv \sigma^{H^+}$ in Table 1, the segment L16-T30
has β-hairpin potential in the aqueous buffer: L16-N21 $\overline{FP}_i = -0.0690$ (C_5 strand), N22-I26 $\overline{FP}_i = 0.6667$ (turn),
and L27-T30 $\overline{FP}_i = -0.0964$ (C_5 strand); (b) The FP_i plot for amylin at low pH and the putative helical

conformation of the peptide. Assuming $\sigma^{His} \equiv \sigma^{H+}$ in Table 1, the segment T9-N22 has unambiguous α-helical potential in the aqueous buffer, $\overline{FP}_i = 0.1746$. **(B) Acid-induced loop-to-helix transition in influenza hemagglutinin** [158]. The trimeric glycoprotein hemagglutinin from influenza virus acts as a fusogen at low pH of endocytic vesicles. The activity is contingent on a large-scale structural rearrangement crucial for delivering the viral contents into host cells. The rearrangement involves inter alia conversion of B-loop of hemagglutinin (residues 55–76) into a long α-helix: (a) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plot for B-loop at high pH. The segment has a helical FP_i profile but very low $\overline{FP}_i$ and it is accordingly unstructured; (b) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plot for B-loop at low pH. The helical data cluster is now shifted into the stable 'helix' region, $\overline{FP}_i > 0$. **(C) Partial unfolding of the translocation-domain helices of the α-pore forming diphteria toxin at low pH** [159]. The toxin invades a cell by crossing the endosome membrane via the process mediated by the C-terminal segment of the all-α domain T (translocation domain) and triggered by the reduced pH in the endosomal lumen: (a) At the standard physiological pH, the toxin is well-structured and the C-terminal segment of the domain T is buried under the helices TH1-TH4, PDB ID 1f0l; (b) At low pH, the TH2, TH3 and TH4 helices become completely disordered, PDB ID 4ow6 [159]. As a result, the C-terminal helices are exposed and can interact with the membrane; (c) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plot for the domain T at the neutral pH. The plot shows the characteristic pattern of a soluble multi-helix bundle, cf. Fig 15A(e), where α-helices incorporate 'C$_5$ strand' segments necessary to stabilize the compact tertiary structure; (d) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plot for the domain T at low pH: the folding potential FP_i becomes more positive for the entire domain (the 'C$_5$ strand' segments are 'titrated out' which destabilizes the tertiary structure); (e)-(h) The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots for the α-helices TH3 and TH4 at the neutral pH, panels (e) and (g), and at the low pH, panels (f) and (h). In both cases the folding potential shifts from $FP_i \lesssim 0$ in panels (e) and (g) (the α-helices which are stable in the interior of a compact structure) to $FP_i \geq 0.3$ in panels (f) and (h) (the α-helices which are unstable except in a highly polarizing environment cf. Fig 7). **(D) Acid-induced unfolding and aggregation of transthyretin** [38,162–172]: (a) The superposed FP_i plots for transthyretin at pH >7 and pH 4, and the DSSP assignments for the native monomer (homotetramer subunit, β sandwich fold). The red and purple outlines mark the strand and turn segments most affected by the reduction of pH: FP_i of the V30-F33 segment (βB strand) shifts from 'C$_5$ strand' to 'helix' propensity, and the two 'turn' segments (D-E≡D* and E-F≡E*) are destabilized. In addition, FP_i of βC and βE segments shifts from 'C$_5$ strand' to 'C$_{7eq}$ strand' propensity. These changes destabilize the outer β sheet of transthyretin βC↓βB↑βE↓βF↑; (b) Putative amyloidogenic monomers of transthyretin generated by step-by-step dismantling and rearrangement of transthyretin β sandwich consistent with the destabilizing FP_i shifts and the solid state NMR [162–164], spin-labelling [165,166], immunoreactivity [167] and H/D exchange (HXMS) [168,169] studies; (c) The pathway of acid-induced unfolding and aggregation of transthyretin [38], and the topology of annular octamers [170–172] and protofibrils [169].

mutations (including G38V) which considerably increase C$_5$ propensity of β3, and increase threefold stability of dimers [217]; (5) C-terminal extensions of the Aβ length that increase C7$_{eq}$/C$_5$ propensity of β3 (e.g. in Aβ$^{1-41}$, Aβ$^{1-42}$, Aβ$^{1-43}$ or Aβ$^{1-46}$ cf. the FP_i plots in the panels A(a)-(d), are reported, albeit with some exceptions [218], to facilitate oligomerization [219–228].

Thus, the first step of aggregation of Aβ is coupled folding and binding resulting in formation of the staggered antiparallel dimer, Fig 22B(c). By analogy to the two-stranded antiparallel β-sheet that anchors the horseshoe-like fold of inteins, cf. section **b**.(ix) and Fig 20, the antiparallel Aβ dimer is expected to fold into a left-handed coiled-coil that has in addition a left-handed superhelical twist. The antiparallel assembly is stabilized by matching the CHC and C-terminal 'C$_5$ strand' segments, and by H-bonding between the segments of consecutive 'turns' which may adopt the C$_5$* fold. The D1-K16 and E22-K28 segments are not involved in β structure of the homodimer, the first remaining 'free', as the unattached 'C$_{7eq}$' strand, and the other forming a loop. The *intermolecular* contacts of the C$_5$* segments implied by the dimer's register are observed in some Aβ aggregates as shown in the inserted diagram [201].

To ensure parallel dimerization of Aβ via *FP*-directed molecular recognition, the β1 and β3 segments would have to adopt the C$_{7eq}$ configuration while the β2 segments adopt the C$_5$* configuration as shown in Fig 22B(d). This mode of assembly yields out-of-register parallel β-sheets expected to fold with ease into parallel β-barrels. The *intermolecular* residue contacts implied by this model, e.g. F19 vs. I31 or L34, are in fact observed in some Aβ oligomers and
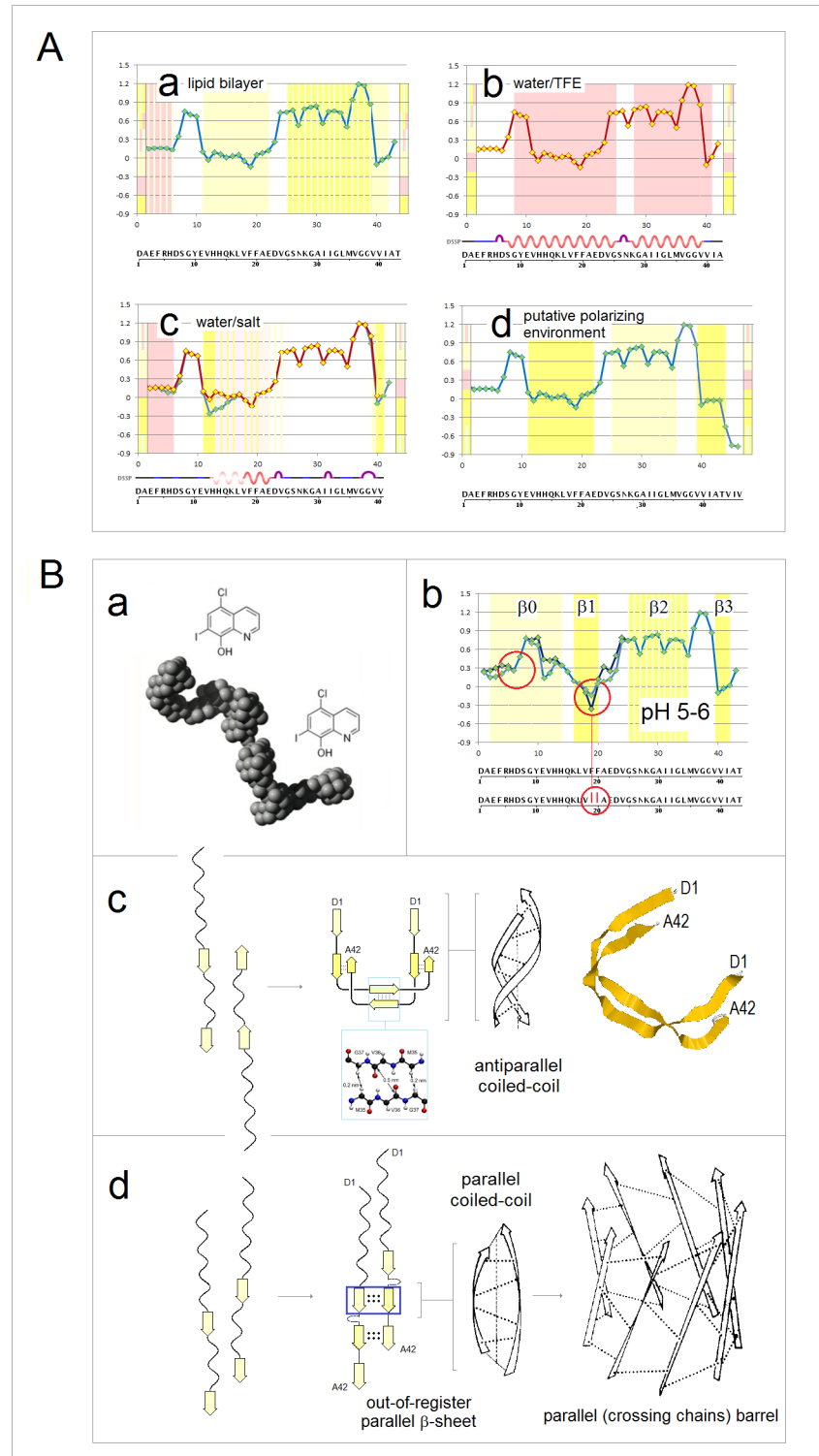
Fig 22. Electronic configuration of the polypeptide backbone and conformational behavior of Aβ proteins. (A) Folding potential, medium effect and secondary structure of monomeric Aβ. The FP_i plots for Aβ peptides are juxtaposed with the color-coded bars representing medium effects cf. Fig 7 (note that at pH 7, σ^{His} is average of the H^0 and H^+ constants since pK_a values of Aβ histidines are in the 6.8-7.0 range [174,175]): (a) **Nonpolar environment**: e.g. bilayer membrane/lipid matrix (FP_i plot for Aβ$^{1-43}$). Elements of the E11-E22 and C-terminal segments are expected to adopt the C$_{7eq}$ fold while the array of consecutive

'turns' D23-M35 may be stabilized as the 'C$_5$* strand'; in contrast, the E3-D7 segment may adopt a helical conformation. The G25-V39 segment of Aβ$^{1-40}$ bound to DMPC bilayer (multilamella vesicles composed of 1,2-dimyristoyl-sn-glycero-3-phosphocholine, high P/L ratio) is indeed reported to fold into a parallel β structure [176,177]. On the other hand, the E3-D7 residues buried in the CDR loops of the complexes with monoclonal antibodies are found to be helical: PP$_{II}$- (PDB ID's 2ipu, 3bae) or 3$_{10}$- (PDB ID 4hix) [178]. (b) **Moderately polarizing environment**: e.g. H$_2$O/TFE (HFIP) solutions or micelle interfaces (*FP$_i$* plot for Aβ$^{1-42}$). The E11-E22 and D23-M35 segments are expected to display α-helix and α*-helix propensities, respectively; the E3-D7 segment is expected to remain disordered since it lacks a well-defined conformational propensity. Several reports confirm that Aβ peptides adopt an all-α fold in the environments of micellar interfaces or aqueous solutions of fluorinated alcohols (see e.g. PDB ID 1iyt) [179–184]. (c) **Polar environment**: e.g. aqueous buffers (red *FP$_i$* plot for Aβ$^{1-40}$ at pH 7; cyan *FP$_i$* plot for Aβ$^{1-43}$ at pH 8–9). In view of the *FP$_i$* patterns, the E11-E22 and D23-M35 segments are expected to remain disordered in a neutral aqueous buffer. However, the E11-E22 segment may sample helical conformations while the CHC (L17-A21), E11-Q15, and C-terminal residues may sample either the C$_5$ (pH 8, longer isoforms) or the C$_{7eq}$ conformations. In agreement with these expectations, experimental evidence suggests that Aβ monomers form in aqueous solution collapsed coils [185–188], devoid of secondary structure [189,190], but the evidence of stabilization of residual secondary structure of Aβ in some conditions is also reported (e.g. PDB ID 2lfm and 2otk) [191–194]. (d) **Highly polarizing environment**: e.g. the environment within an amyloid fibril (*FP$_i$* plot for Aβ$^{1-46}$). The E11-E22 and D23-M35 segments are now expected to adopt C$_5$ and C$_{7eq}$ conformations respectively. The distribution of amyloid-fiber forming capacities, cf. Fig 11C, is consistent with this expectation. **(B) Electronic Configuration of the Polypeptide Backbone and Aggregation of Aβ**. (a) The Aβ$^{1-42}$ dimer obtained by modelling studies based on the SAXS data for the stoichiometric Aβ$^{1-42}$ complexes with 8-hydroxyquinolines [200]; (b) The *FP$_i$* plot shows the anticipated 'strand' configurations in the Aβ homodimer at pH 5-6 i.e. under the conditions that accelerate fibrillization (light blue curve H$^+$/E$^-$, blue curve H$^+$/E^0) [203,204]; (c) The antiparallel coiled coil as a model of Aβ dimer structure: the favoured antiparallel dimerization of Aβ in aqueous solutions matches the CHC and C-terminal segments that both adopt either the C$_{7eq}$ or C$_5$ conformations. The antiparallel assembly could in addition be stabilized by H-bonding between the segments of consecutive 'turns' V24-G37 which may adopt the C$_5$* fold in the dimer. The *intermolecular* residue contacts implied by this model are observed in some Aβ aggregates as shown in the insert [201]. By analogy to the two-stranded antiparallel β-sheet of inteins, cf. Fig 20, the antiparallel Aβ dimer is expected to fold into a left-handed coiled-coil, shown in the diagram with a left-handed superhelical twist; (d) The parallel coiled coil as a model of Aβ dimer structure: the favoured parallel dimerization of Aβ in aqueous solutions would match the CHC and 'turns' segments, the C$_{7eq}$ and C$_5$* folds respectively, in both combinations. This mode of assembly yields out-of-register parallel β-sheets expected to fold with ease into parallel β-barrels. The *intermolecular* residue contacts implied by this model, e.g. F19 vs. I31 or L34, are in fact observed in some Aβ oligomers and aggregates [202]. However, neither parallel dimer nor parallel barrel are likely to form and persist in absence of a stabilizing molecular 'template'.

aggregates [202]. However, neither parallel dimer nor parallel barrel are likely to form and persist in absence of a stabilizing molecular 'template'.

(ii) A model for *in vitro* dimerization of Aβ peptides: Effects of concentration, small-molecule modulators, and surface/interface support. Formation of the dimer described in Fig 22B is contingent on the 'unfolding' of the collapsed-coil ensemble of Aβ. This transition will be facilitated by transferring the protein from an aqueous solution to a less polar environment which drives the N-terminal half of Aβ into a helical conformation, cf. Fig 22A(b). Such a transfer can be achieved in several ways, see Fig 23:

1. If Aβ concentration in an aqueous buffer is sufficiently high, c>c*, the protein initially forms micelle-like aggregates, Fig 23A [229–231]. The protein is thus buried in a low-permittivity folding basin *FB* and its N-terminal half adopts helical conformation which 'unwinds' the collapsed coil. Experimental evidence suggests that the α-helical Aβ conformer is indeed an obligatory intermediate under these conditions [232,233]. Since the D23-K28 segment forms a side loop in the two-stranded staggered antiparallel β-sheet, stabilization of this loop by a lactam or a disulfide bridge will also facilitate dimerization [234,235].

2. If Aβ concentration is too low to generate pseudomicellar aggregates, c<c*, dimerization of Aβ may be catalysed by the detergents and detergent-like amphipathic co-solutes [236–
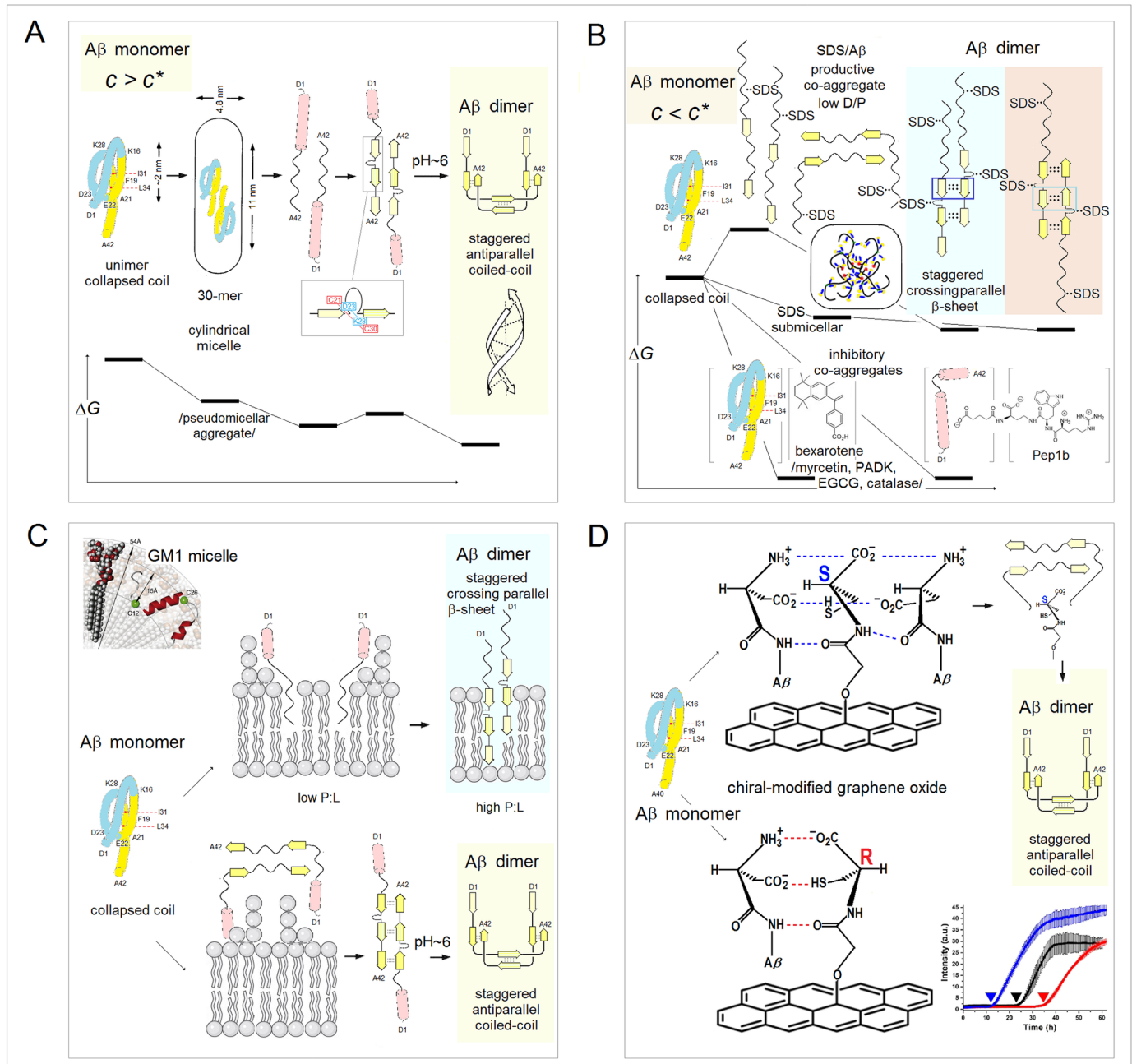
Fig 23. A model for *in vitro* dimerization of Aβ peptides: Effects of concentration, small-molecule modulators, and surface/interface support. Formation of the dimer described in Fig 22B is contingent on the 'unfolding' of the collapsed-coil ensemble of Aβ. This transition will be facilitated by transferring the protein from an aqueous solution to a less polar environment which drives the N-terminal half of Aβ into a helical conformation. Such a transfer can be achieved in several ways: **(A)** If Aβ concentration in an aqueous buffer is sufficiently high, c>c*, the protein initially forms micelle-like aggregates [229–231]. The protein is thus buried in a low-permittivity folding basin and its N-terminal half adopts a helical conformation which 'unwinds' the collapsed coil. **(B)** If Aβ concentration is too low to generate pseudomicellar aggregates, c<c*, dimerization may be catalysed by the detergents and detergent-like amphipathic co-solutes [236–238]. E.g. SDS at submicellar concentrations is shown to form Aβ co-aggregates [239] in which protein's N-termini are assembled in the center of the aggregate while the C-terminal chain segments protrude into solution. On the other hand, however, small-molecule co-solutes may also stabilize the collapsed-coil ensemble or the helical conformer of Aβ and thereby prevent dimerization [240–246]. **(C)** Binding of Aβ in a water/micelle or water/lipid membrane interface transfers the protein from an aqueous buffer into a less polar environment. The expected conformational transition is well documented, cf. Fig 22A(a) and 22A(b) [176,177,181–184], and may well contribute to the catalysis of polymerization by membranes and lipid rafts [247–254]. **(D)** (a) The (*S*)-cysteine can bring together two Aβ molecules via three-point binding of Asp-1, in spite of a degree of steric strain in either chelation mode; (b) The (*R*)-cysteine can tightly bind one Aβ molecule in a complex free of strain, but bringing in the second Aβ molecule is impossible due to prohibitive steric hindrance; (c) Thus, the (*S*)-cysteine would

promote formation and release of the Aβ↑Aβ↓ dimer and therefore accelerate polymerization (the blue curve in the insert), while the (R)-cysteine would immobilize isolated Aβ molecules on the graphene surface and therefore slow down polymerization (the red curve in the insert).

238]. E.g. SDS at submicellar concentrations is shown to form Aβ co-aggregates [239] in which protein's N-termini are assembled in the center of the aggregate while the C-terminal chain segments protrude into solution and may form either parallel or, more likely, antiparallel β-sheets, Fig 23B. However, small-molecule co-solutes may also stabilize the collapsed-coil ensemble or the helical conformer of Aβ and thereby prevent dimerization [240–246].

3. Binding of Aβ in a water/micelle or water/lipid membrane interface transfers the protein from an aqueous buffer into a less polar environment, Fig 23C. The expected conformational transition is well documented, cf. Fig 22A(a) and 22A(b) [176,177,181–184], and may well contribute to the catalysis of polymerization by membranes and lipid rafts [247–254].

Binding of Aβ on the cysteine-modified graphene oxide surface either accelerates or slows down fibrillization depending on the chirality of cysteine [40]. The proposed model of Aβ dimerization suggests that this effect can be explained by the difference in how the N-terminal Asp residue (Asp-1) of 'unfolded' Aβ (in water/graphene interface) interacts with the (S)- and (R)-cysteine, marked in blue and red respectively in Fig 23D. The (S)-cysteine can bring together two Aβ molecules via three-point binding of Asp-1, in spite of a degree of steric strain in either chelation mode, Fig 23D(a). On the other hand, the (R)-cysteine can tightly bind one Aβ molecule in a complex free of strain, but bringing in a second Aβ molecule is impossible due to prohibitive steric hindrance, Fig 23D(b). Thus, the (S)-cysteine would promote both formation and release of the Aβ↑Aβ↓ dimer and therefore accelerate polymerization (the blue curve in the insert, Fig 23D(c)), while the (R)-cysteine would immobilize isolated Aβ molecules on the graphene surface and therefore slow down polymerization (the red curve in the insert).

(iii) Aggregation of Aβ dimers into paranuclei by domain swapping.   Several lines of evidence support the notion of independent pathways of oligomerization and fibrillization of Aβ proteins [237,255–262]. The staggered antiparallel structure implies that the initial association of two Aβ dimers can proceed either (1) by domain swapping between twisted coiled-coil conformers i.e. antiparallel interlocking of the 'free' N-terminal segments, Fig 24A(b), or (2) by the edge-to-edge assembly of the extended conformers, Fig 24B(b). The difference in conformational behavior of the resulting tetramers gives rise to those independent pathways of polymerization: domain swapping leads to parallel cross-β structure, Fig 24A(g), while edge-to-edge assembly leads to antiperiplanar β barrel and antiparallel cross-β structure, Fig 24B(c) and 24B(d).

The tetramers obtained by domain swapping are stabilised by the increased C_{7eq} propensity of these segments and by the terminal salt-bridges hence the domain-swapping pathway will be favoured at low pH and low salt; these conditions were in fact shown to accelerate fibrillization [202,203,263,269]. Further aggregation via domain swapping yields a disk-shaped hexamer—paranucleus [204,264–268], see Fig 24A(c). The key feature of this circular hexamer is the superposition of two out of the three antiparallel coiled-coil dimers which produces a wedge-like shape and the characteristic trapezoid appearance in the high-resolution AFM images [266], see the insert in Fig 24A(c). The superposition places the C_5^* strands, viz β/β′ and β″/β‴ strands in the β structure diagram in Fig 24A(c), on top of each other in the parallel alignment so that the paranucleus is stabilized by the 'parallel steric zipper' interface.
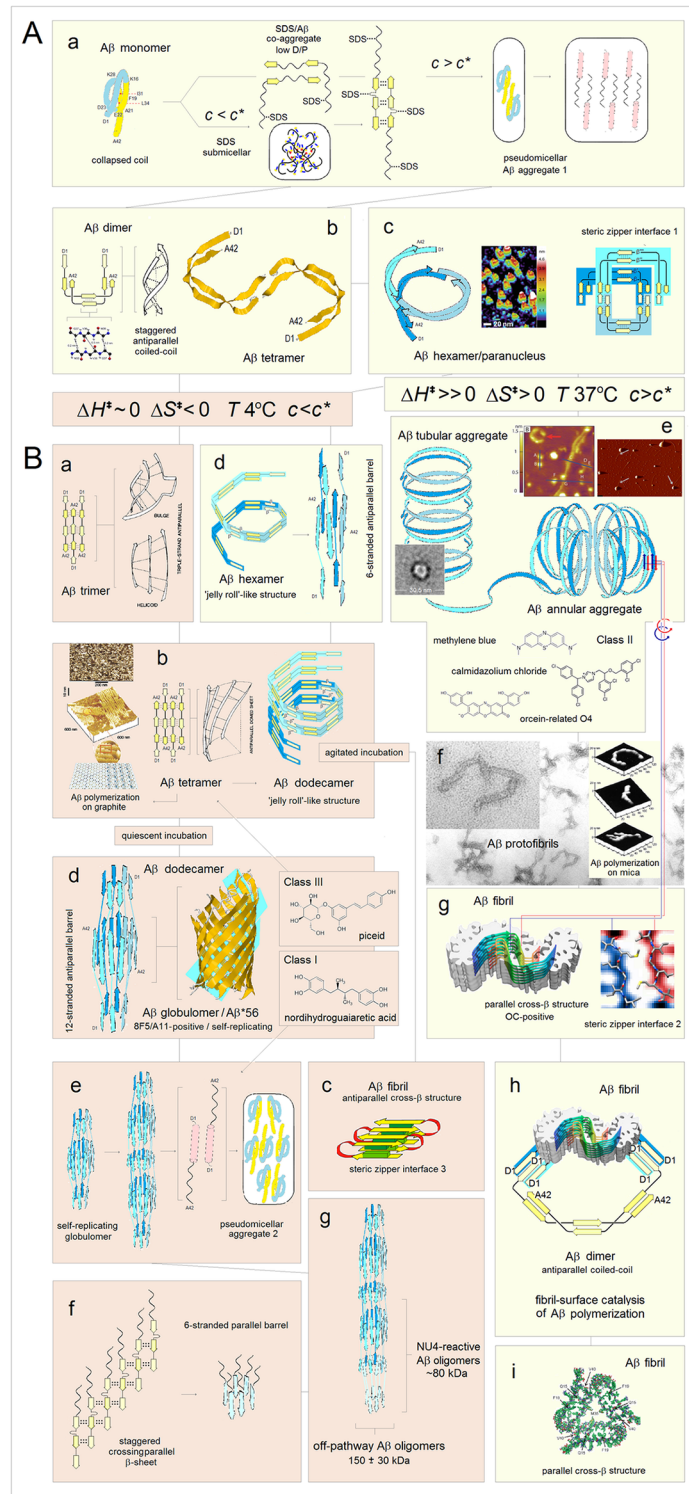
Fig 24. A model for nucleated polymerization of Alzheimer's Aβ proteins. (A) Domain swapping: Paranuclei and fibrils comprising parallel β-sheets. (a) The mechanism of conversion of the Aβ collapsed coils into out-of-register antiparallel dimers via split intein-like mechanism of molecular recognition, (1) at the Aβ concentration greater than the critical micelle concentration $c > c^*$, and (2) at the Aβ concentration $c < c^*$, in the presence of submicellar concentrations of SDS; (b) The staggered antiparallel Aβ dimer as the left-handed coiled-coil that has in addition left-handed superhelical twist, and the Aβ tetramer obtained via

antiparallel association of the 'free' N-terminal segments (domain swapping); (c) Further aggregation via domain swapping yields a disk-shaped hexamer—paranucleus [204,264–268]. The circular complex places two dimers on top of each other. The characteristic trapezoid appearance of paranuclei in high-resolution AFM images [266], see the insert, appears to be consistent with the wedge-like shape. The strands stacked on top of each other are parallel; (d) The circular hexamer complex may also be stabilized by the edge-to-edge H-bonding within a 'jelly-roll'-like structure which may convert into a 6-stranded antiparallel β barrel; (e) The limiting modes of stacking of the paranuclei which yield either tubular (Aβ amyloid pore [271]) or annular aggregates (Aβ nanoglobules [272]). The high-resolution AFM image of the initial stages of aggregation of the paranuclei shows stacking of disk-shaped hexamers and formation of a ring structure (red arrow). Catalysis of fibrillization by the intercalating cations and anions e.g. methylene blue, calmidazolium chloride, orcein-related O4 [276–278] suggests that conformational conversion of the antiparallel sheets β↑β'↓ and β''↑β'''↓ of the paranuclei into parallel cross-β structure occurs in such stacks; (f) Morphological diversity of polymerization on mica support [42] appears to reflect two modes of paranuclei aggregation, and so does the structure of protofibril 'on-path' intermediates of fibrillization in solution [279–282]; (g) The cryoEM-derived structures of Aβ fibrils shows the expected assembly of two protofilaments comprising parallel cross-β sheets and aligned in the antiparallel fashion [283,284]; (h) Mechanism of fibril-surface catalysis of secondary nucleation [43,285,286]: formation of the Aβ dimers is facilitated by the binding of Aβ monomers along the fibril edges via antiparallel interlocking of the N-terminal segments; (i) The trimeric fibril which may be formed by remodeling of the fibril shown in panel (g) [288]. **(B) Edge-to-edge assembly: Off-pathway oligomers, self-replicating non-fibrillar aggregates and fibrils comprising antiparallel β-sheets**. (a) The Aβ trimer obtained by the edge-to-edge assembly. In contrast to the dimer, the trimer cannot fold into a coiled coil [154]; (b) Morphology of polymerization of Aβ on graphite [41] is consistent with the model of edge-to-edge assembly and *FP*-directed molecular recognition, see the text. The edge-to-edge Aβ tetramer cannot fold into a coiled coil either [154] but its polymerization via domain swapping may produce high-order oligomers that retain the staggered antiparallel alignment of strands [261,289–291]. By a combination of twist, bend and rise, such oligomers can fold into 'jelly-roll'-like cylindrical structures; (c) Reversal of domain swapping and fragmentation within the 'jelly-roll'-like folds of higher-order oligomers can yield fibrils comprising antiparallel cross-β structure [292,293]; (d) The alternative conversion of the 'jelly-roll'-like cylinder involves reversal of domain swapping and the edge-to-edge closing of a β barrel. The homotrimer of tetramers can form in this way the 12-stranded antiparallel barrel which we believe represents the structure of the neuropathological Aβ globulomer [258,294] also isolated as the brain Aβ*56 oligomer [262,295,296]; (e) The 12-stranded antiparallel barrel can 'capture' Aβ dimers and tetramers via antiparallel interlocking of the 'free' N-terminal segments. This is likely the mechanism of self-replication of the off-pathway Aβ oligomers as shown in the diagram [260,297–299]. Note that the low-permittivity environment of the interior of a larger aggregate may promote strand→helix conversion in reversal of the initial stages of dimerization of Aβ [301]; (f) The alternative edge-to-edge aggregation of Aβ, the staggered (out-of-register) parallel assembly; (g) The 12-stranded antiparallel barrel may also 'capture' Aβ monomers that form a staggered parallel assembly and fold into 6-stranded out-of-register parallel barrel. This type of aggregation would account for the formation of the off-pathway Aβ oligomers 150±30 kDa [201,202] and the NU4-reactive Aβ oligomer ~80 kDa [300] (30-mer and 18-mer, respectively), and for the solid state NMR evidence of *both antiparallel and parallel intermolecular* contacts in the 150±30 kDa oligomers [201,202].

Consequently, a rotation about the strand axis would be sufficient to convert the antiparallel sheets β↑β'↓ and β''↑β'''↓ into parallel sheets β↑β''↑ and β'↓β'''↓. This conformational conversion would encounter a considerable activation energy barrier in spite of the positive entropy of activation: $\Delta H^{\ddagger} >> 0$, $\Delta S^{\ddagger} > 0$ [202,203,269,270].

(iv) Aggregation of paranuclei, conversion of antiparallel β structure into parallel cross-β-structure, and fibril-surface catalysis of secondary nucleation of Aβ. There are, given the wedge-like profile of the disk-shaped paranucleus, two limiting modes of self-aggregation of paranuclei that can be stabilized by the 'steric zipper' interfaces: the annular stacking —all paranuclei 'wedge out' in the same direction, and the tubular stacking—the neighboring paranuclei 'wedge out' in opposite directions. These limiting modes of stacking of the paranuclei yield either tubular (Aβ amyloid pore [271]) or annular aggregates (Aβ nanoglobules [272]). The high-resolution AFM image of the initial stages of aggregation of the paranuclei, shown in the central insert in Fig 24A(e), is consistent with this notion: 'Features A, B, and C indicate Aβ hexamers, and D, E, and F indicate Aβ dodecamers consisting of two stacked hexamers', and the red arrow points to an annular aggregate [273]. The morphology of Aβ polymerization on mica [42], where the paranuclei would be held on the hydrophilic surface via

backbone H-bonding i.e. bound upright, is also consistent with the two modes of paranuclei stacking: the aggregates form either large ring structures or elongated bar structures, see the inserts in Fig 24A(f).

Formation of a stack facilitates conformational conversion of the antiparallel sheets β↑β′↓ and β″↑β‴↓ of the paranuclei into parallel cross-β structure [272] which is stabilized within the pre-organized architecture of the aggregates [274,275]. This process is also facilitated by the intercalation of the aromatic cations or anions e.g. methylene blue [276], calmidazolium chloride [277] or orcein-related O4 [278], see Fig 24A(e), and the relatively low H-bonding propensity of the 'C$_5$* strands' i.e. the least polarized backbone segments.

Morphological diversity of the protofibril 'on-path' intermediates of fibrillization in solution [279–281] suggests that both paths of paranuclei assembly are also followed in the intermediate stages of polymerization [282], Fig 24A(f). However, the annular aggregates of paranuclei might more readily undergo the conformational conversion. Such a conversion would yield fibrils consisting of two protofilaments that comprise parallel cross-β sheets in antiparallel alignment and this arrangement is indeed observed in the cryoEM-derived structures of Aβ fibrils [283,284], Fig 24A(g). These fibrils can catalyze formation of the staggered antiparallel dimers by 'unwinding' collapsed-coil Aβ monomers via antiparallel interlocking of the N-terminal segments (domain swapping) along the fibril's edges and by ensuring monomers' proximity [43,285,286], Fig 24A(h). Lastly, folding of the cross-β sheets shown in Fig 24A(g) into β arcades may initiate separation of the two protofilaments into monomeric fibrils [287], and remodeling of those fibrils into the dimeric and trimeric fibrils [288], Fig 24A(i).

(v) Edge-to-edge assembly: Off-pathway oligomers, self-replicating non-fibrillar aggregates and fibrils comprising antiparallel β-sheets. The alternative pathway of aggregation of Aβ proteins [236–238,255–264,289–301] begins, we propose, as the edge-to-edge complexation of extended dimers and monomers to form staggered antiparallel trimers, Fig 24B(a) (the trimer previously identified as a highly toxic oligomer [259] may have this structure), or tetramers, Fig 24B(b). The morphology of polymerization of Aβ on graphite results from this type of assembly: the extended Aβ chains are held on graphite by the sidechains' interactions with the hydrophobic surface and consequently assemble edge-to-edge via backbone H-bonding. The long antiparallel monolayer stacks of Aβ are ~19 nm wide [42] i.e. the chains are staggered (extended peptide is ~15 nm long). Thus, these stacks can catalyze alongside polymerization via domain swapping; this could be the origin of the characteristic clusters of parallel stacks on graphite, see the inserts in Fig 24B(b).

The oligomers obtained by the edge-to-edge assembly cannot fold into coiled coils [154] but further polymerization via domain swapping may produce high-order oligomers that retain the staggered antiparallel alignment of strands [261,289–291]. By a combination of twist, bend and rise, such oligomers can fold into 'jelly-roll'-like cylindrical structures. Fragmentation within the cylindrical folds of higher-order oligomers can yield fibrils comprising antiparallel cross-β structure [292,293], Fig 24B(c). The alternative conversion of the 'jelly-roll'-like cylinder may involve the edge-to-edge closing of a β barrel and a concomitant reversal of domain swapping. The homotrimer of tetramers can form in this way the 12-stranded antiparallel barrel which we believe represents the structure of the neuropathological Aβ globulomer [258,294], also isolated as the brain Aβ*56 oligomer [262,295,296], Fig 24B(d). Class III assembly modulators apparently promote the dissociation of such barrels while Class I modulators promote the formation of pseudomicellar aggregates [276].

The proposed structure suggests a mechanism of self-replication of the off-pathway oligomers [260,297–299]. The N-terminal segments of Aβ, which are protruding from the 12-stranded antiparallel barrel, are not expected to form stable in-register 6-stranded parallel barrels. Thus, these N-terminal segments can 'capture' the staggered antiparallel Aβ dimers or

tetramers via domain swapping and catalyze in this way the assembly of the identical 12-stranded barrel, Fig 24B(e).

In principle, the edge-to-edge aggregation of Aβ may also result in the staggered (out-of-register) parallel assembly, Fig 24B(f). This type of β structure will have a tendency to form a parallel β barrel, cf. Fig 22B(d), as long as it is stabilized e.g. by the antiparallel barrel serving as a molecular 'template' as shown in Fig 24B(g). Here, the N-terminal segments on one side of the 12-stranded barrel bind just 6 Aβ monomers which can then form stable 6-stranded out-of-register parallel barrel. This type of aggregation would account for the formation of the off-pathway Aβ oligomers 150±30 kDa [201,202] and the NU4-reactive Aβ oligomer ~80 kDa [300] (30-mer and 18-mer, respectively), and for the presence of *both antiparallel and parallel intermolecular* residue contacts in the 150±30 kDa oligomers [201,202]. Note that the low-permittivity environment of the interior of such large aggregates may promote strand→helix conversion in a reversal of the initial stages of dimerization of Aβ [301].

(vi) Morphology of Aβ aggregation on the NIBC monolayers on gold: Combined effects of cysteine chirality and wedge-like shape of Aβ paranuclei.   The notion of the wedge-shape control of aggregation morphology may explain the effect of chirality on the assembly of Aβ(1–40) on the self-assembled monolayers of *R(S)*-N-isobutyrylcysteine (NIBC) on gold [41]. By analogy to Aβ polymerization on mica, section **b**.(iv) and Fig 24A(f), aggregation of Aβ(1–40) on NIBC monolayers produces large ring structures and elongated bar structures. However, the course of aggregation depends on the configuration of the amino acid: large rings form on *L*-NIBC monolayers, and elongated bars form on *D*-NIBC monolayers, see the inserts in Fig 25. Binding of the H14-K16 segment of Aβ to the cysteine monolayer, Fig 25(a), was shown to be crucial to observe this effect [41]. Given that the NIBC monolayers comprise pairs of long files of N-isobutyrylcysteines [302], Fig 25(b), we assume that the putative complex of the Aβ paranucleus on the monolayer surface involves backbone H-bonding of the extended H14-K16 segment to the isobutyryl carbonyls of one file of N-isobutyrylcysteines in the pair (i.e. paranucleus is bound upright). The positively charged side chains of His-14 and Lys-16 extend on the same side of the strand and can stabilize the complex by binding to the isobutyryl carbonyls of the second file of N-isobutyrylcysteines in the pair. On the surface of *L*-NIBC monolayer [302], binding of the H14 side chain to the isobutyryl C = O of the neighbouring row of cysteines is unimpeded; the K16 side chain encounters an impediment but its length and flexibility make the binding possible, see Fig 25(c). Thus, the tight packing of paranuclei is achieved via parallel alignment of the H14-K16 segments which directs all the complexed paranuclei to 'wedge out' in one direction and promotes annular stacking and formation of large ring structures. In contrast, on the surface of *D*-NIBC monolayer [302], H14 side chain cannot bind to the isobutyryl C = O of the neighbouring row of cysteines: it encounters an impediment and lacks the flexibility and length needed to overcome the hindrance, Fig 25(d). Thus, the tight packing on this surface is achieved via antiparallel alignment of the H14-K16 segments which directs the neighbouring paranuclei to 'wedge out' in opposite directions and thereby promotes tubular stacking and formation of elongated bar structures.

(vii) Fibrillization of tau protein and α-synuclein via antiparallel coiled-coil pathway. Tau is a highly soluble, intrinsically disordered protein [303,304]. Its sequence includes the 'repeat domain' which binds to lipid membrane surfaces [305,306] and microtubules [307,308] via short helices, Fig 26A(a), folds into β-hairpins upon binding to β-wrapin TP4 [309], and forms amyloid fibrils. The FP_i profile of the 'repeat domain' construct K19 (R1-R3R4) at pH 7 shows Aβ-like pattern of backbone polarization which corresponds to the sequence C_{7eq}-C_5-C_5^*-C_5 of anticipated 'strand' configurations, Fig 26A(b); the FP_i-based assignment of secondary structure is consistent with the solid-state NMR data on the K19 amyloid fibrils [310]. The FP_i plot also shows that the 'C_5 strand' propensity of the V337-E342 segment (VEVKSE) is lost
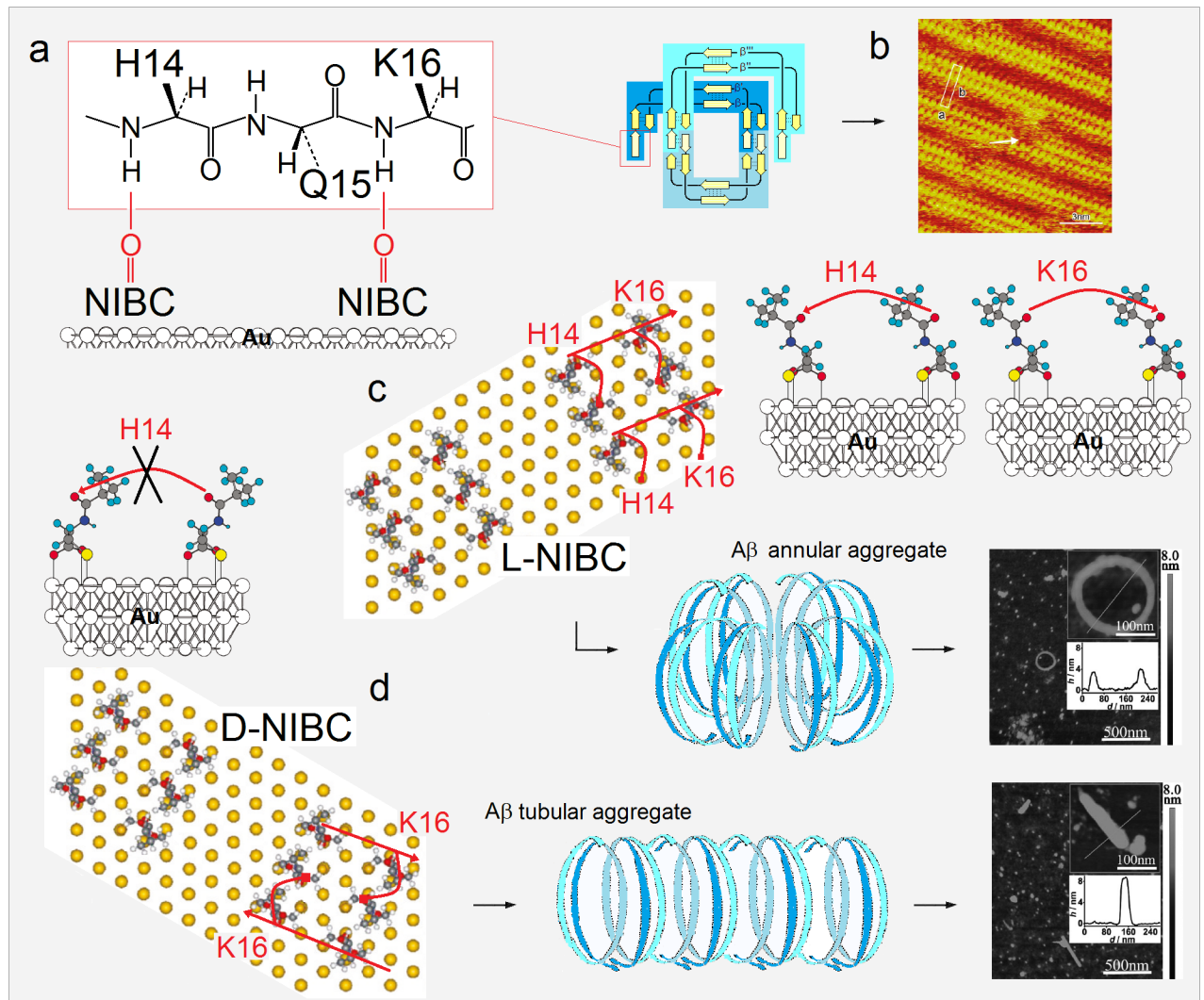
Fig 25. Morphology of Aβ aggregation on the NIBC monolayers on gold: The combined effects of cysteine chirality and wedge-like shape of Aβ paranuclei. (a) The putative complex of the Aβ paranucleus on the monolayer surface involves backbone H-bonding of the extended H14-K16 segment to the isobutyryl carbonyls of N-isobutyrylcysteines [41]. The positively charged side chains of His-14 and Lys-16 extend in the same direction. **(b)** The topology of the self-assembled cysteine monolayer on gold [302]: The monolayer comprises pairs of long files of N-isobutyrylcysteines. **(c)** A model of the surface of L-NIBC monolayer [302], H14 and K16 side chain interactions with the L-NIBC isobutyryl carbonyls, and the putative tight packing of two Aβ paranuclei. The binding of the H14 side chain to the isobutyryl C = O of the neighbouring row of cysteines is unimpeded; the K16 side chain encounters an impediment but its length and flexibility make the binding possible. Thus, the tight packing of paranuclei is achieved via parallel alignment of the H14-K16 segments which directs all the complexed paranuclei to 'wedge out' in one direction i.e. promotes annular stacking and formation of large ring structures. **(d)** A model of the surface of D-NIBC monolayer [302], H14 and K16 side chain interactions with the D-NIBC isobutyryl carbonyls, and the putative tight packing of two Aβ paranuclei. Here the H14 side chain cannot bind to the isobutyryl C = O of the neighbouring row of cysteines; it encounters an impediment and lacks flexibility and length needed to overcome the hindrance. Thus, the tight packing on this surface is achieved via antiparallel alignment of the H14-K16 segments which directs the neighbouring paranuclei to 'wedge out' in opposite directions and thereby promotes tubular stacking and formation of elongated bar structures.

at pH 2, and indeed K19 does not forms fibrils at low pH [310]. Thus, nucleated polymerization, assumed to be the main pathway of fibrillization of tau [311–314], may involve the staggered antiparallel dimers and disk-shaped hexameric intermediate analogical to the paranucleus of Aβ. The annular assembly of the disk-shaped paranuclei and the subsequent conformational conversion (antiparallel→parallel β-structure) would yield fibrils consisting of
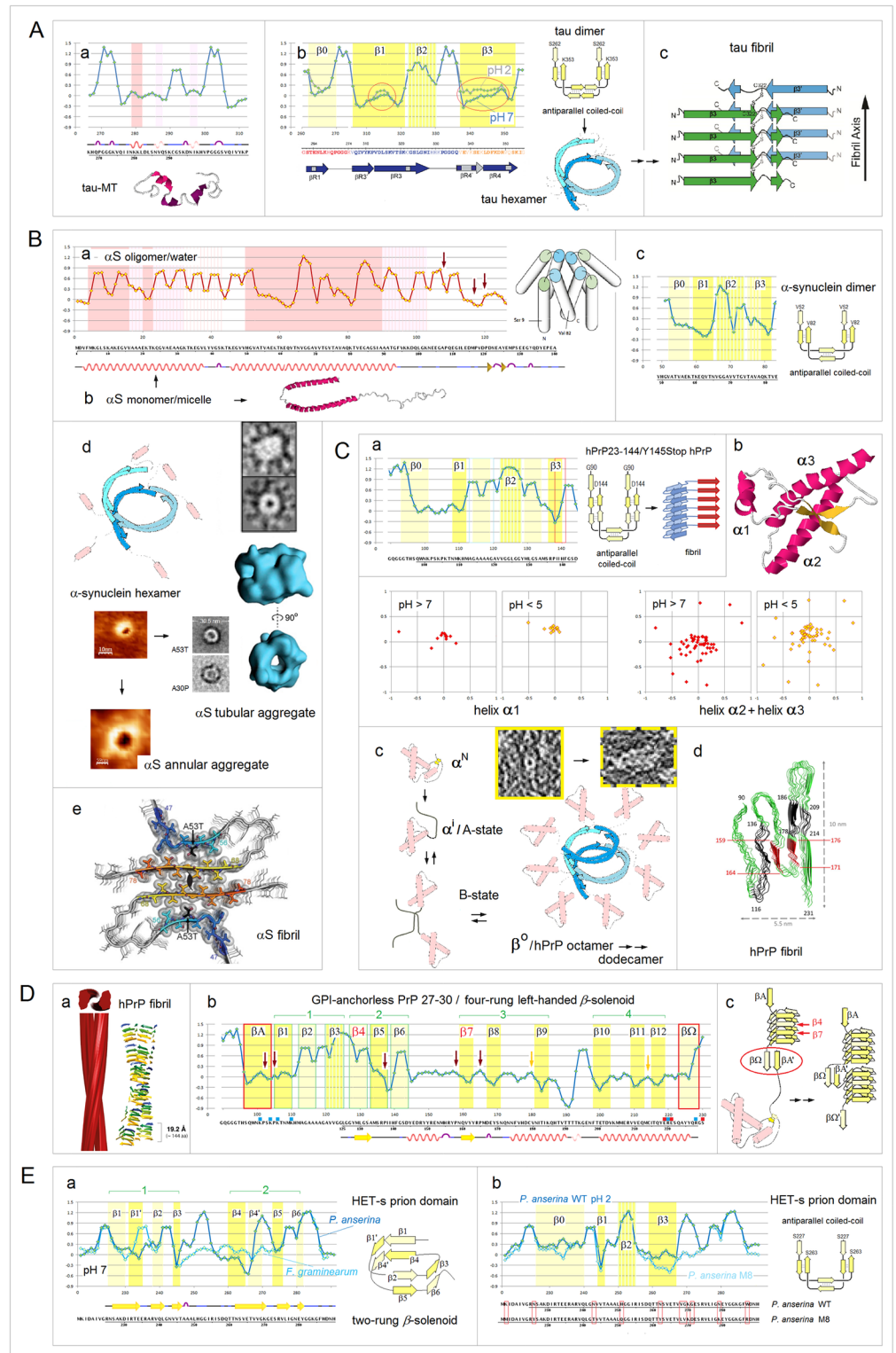
Fig 26. Electronic configuration of the polypeptide backbone and molecular recognition in formation of β structure: Mechanism of aggregation of tau, α-synuclein, human prion hPrP and HET-s PFD. (A) Tau proteins: (a) The FP_i plot and secondary structure of the 46-residue fragment of Tau(267–312) bound to microtubules PDB ID 2mz7 [307]; (b) The FP_i profile of the repeat domain construct K19 (R1-R3R4) at pH 7 is similar to the FP_i profile of Aβ at pH 5. The FP_i-based assignment of secondary structure is consistent with the

solid-state NMR data on the K19 amyloid fibrils [310]. Note that at pH 5 the V337-E342 segment (VEVKSE) has α-helix propensity rather than C_5 propensity required to stabilize the coiled-coil dimer, and K19 actually does not aggregate at low pH [310]. Nucleated polymerization, assumed to be the main pathway of fibrillization of tau [311–314], may involve disk-shaped hexameric intermediate analogical to the paranucleus of Aβ; (c) The annular assembly of the disk-shaped paranuclei and the subsequent conformational conversion (antiparallel→parallel β-structure) would yield fibrils consisting of two protofilaments which comprise parallel cross-β sheets in the antiparallel alignment. The solid-state NMR structure of β-sheet core of tau paired helical filaments confirms the expected arrangement [315]. **(B) α-Synuclein**: (a) The FP_i plot shows secondary structure assignment for a putative tetramer (colour shading reflects helix propensity based on solution NMR data) [318,319], but the existence of such a tetramer is questioned [320,321]; (b) The micelle-bound monomer is helical PDB ID 1xq8 [322,323]; (c) The FP_i profile of the ~35–40 residue-long amyloidogenic region of human α-synuclein which incorporates the NAC sequence essential to fibrillization of this protein. The pattern of backbone polarization corresponds to the C_{7eq}-C_5-C_5*-C_5 sequence of the anticipated 'strand' configurations. Thus, like Aβ and tau, α-synuclein is expected to form the antiparallel, out-of-register, coiled-coil dimer shown in the diagram; (d) The tubular (amyloid pore) [271,324,325] and annular (cf. the insert, diameter ~ 120 nm) [324,326] aggregates of α-synuclein may be formed by the stacking of the wedge-like disk-shaped hexamers analogical to the paranucleus of Aβ. Persistent oligomers of MW ~80 kDa and circular appearance (cf. the insert, diameter ~10 nm) are indeed observed during the polymerization in the presence of heme [326]; (e) Conformational conversion (antiparallel→parallel β structure) within the annular stacks of the paranucleus-like hexamers would yield fibrils consisting of two protofilaments which comprise parallel cross-β sheets in the antiparallel alignment. A recently proposed model of α-synuclein protofibril is consistent with this expectation [332]. **(C) Prion proteins (PrP), conversion to parallel cross-β structure**: (a) The FP_i plot for the recombinant polypeptide hPrP23-144 (a model for Y145Stop variant of human prion, the mutation eliminates the entire α-helical region of hPrP) which undergoes a spontaneous conversion from a monomeric disordered state to the in-register parallel cross-β fibrillar form [335–337]. Ignoring the palindromic segment, see the text, the anticipated β structure configuration of the amyloidogenic region is the sequence C_{7eq}-C_5-C_5*-C_5 (β0-β1-β2-β3) by analogy to Aβ, tau, and α-synuclein. Thus, this region may also form the antiparallel, out-of-register, coiled-coil dimers which assemble via domain swapping into superhelical oligomers and yield fibrils comprising sheets of parallel cross-β structure; (b) The FP_i vs. $\Delta FP_{i-1 \rightarrow i+1}$ plots for the native helices of hPrP at high and low pH. The plots suggest that the helix α1 is destabilized by reduction of pH while the helices α2 and α3 are becoming stable as the compact structure unravels at low pH; (c) The putative pathway of assembly and structure of the octameric paranucleus obtained at low pH [338–348]. The insert shows hPrP aggregates [349] which could be formed by tubular stacking of such octamers or similar higher-order oligomers; (d) The single-sheet parallel in-register cross-β structure of PrP [350,351]. **(D) Prion proteins (PrP), conversion to β-solenoid fibrils**. (a) The structure of fibrils of GPI-anchorless PrP 27–30 [352,353]. The protofilaments are formed by stacking four-rung left-handed solenoids that include nearly entire length of the protein; (b) The FP_i plot for hPrP (omitting most of the N-terminal domain) and the proposed four-rung solenoid structure. The assignment of the strands marked by the green outlines follows the two-rung model [354] proposed as a revision of the original model of the left-handed β helix [355]; the strands β7 through β12 are assigned based on the FP_i minima indicating the 'C_5 strand' propensity. The proposed structure places Cys179 and Cys214 in register (orange arrows) and all Pro residues (brown arrows) in turns and loops. This structure is also consistent with the anticipated β structure configuration of the amyloidogenic region which can be described as the sequence of two 'asymmetrically' complementary arrays of alternant incipient strands: C_5-C_{7eq}-C_5* followed by C_{7eq}-C_5-C_{7eq} (β1-β2-β3 followed by β4-β5-β6 i.e. including the palindromic segment); (c) The model of self-propagation of the infectious prion: the head-to-tail domain swapping between the native conformer and the β solenoid conformer that involves interlocking of the βA and βΩ strands (marked by red outlines in FP_i plot) via assembly of the parallel β sheet $C_{7eq}\downarrow C_5\downarrow$. **(E) HET-s prion forming domain**. (a) The superposed FP_i plots for HET-s prion forming domains from *P. anserina* and *F. graminearum*, and the two-rung left-handed β solenoid fold that the domain from *P. anserina* adopts in amyloid fibrils [356–358]. The anticipated β structure configuration of the *P. anserina* domain is the sequence of two 'asymmetrically' complementary arrays of alternant incipient strands: C_{7eq}-C_5-C_{7eq}-C_5 followed by C_5-C_{7eq}-C_5-C_{7eq} (β1-β1'-β2-β3 followed by β4-β4'-β5-β6); (b) The superposed FP_i plots for the HET-s prion forming domain from *P. anserina* at low pH which induces formation of a generic amyloid fold [363–367], and the M8 variant of this domain that forms toxic amyloids comprising antiparallel β structure [368–370].

two protofilaments which comprise parallel cross-β sheets in the antiparallel alignment. The solid-state NMR structure of β-sheet core of tau paired helical filaments confirms the expected arrangement, Fig 26A(c) [315]; fracture, reassembly and growth of those fibrils may be the source of polymorphism [316].

α-Synuclein is a highly pleiomorphic protein as well [317–323], Fig 26B. The FP_i plot for the ~35-40 residue-long amyloidogenic region of human α-synuclein which incorporates the

NAC sequence essential to fibrillization of this protein, is shown in Fig 26B(c). Again, the pattern of backbone polarization corresponds to the familiar sequence of the anticipated 'strand' configurations C_{7eq}-C_5-C_5^*-C_5. Thus, like Aβ and tau, α-synuclein is expected to form the antiparallel, out-of-register, coiled-coil dimer. The tubular (amyloid pore) [271,324,325] and annular (cf. the insert, diameter ~120 nm) [324,326] aggregates of α-synuclein may be formed by the stacking of the wedge-like disk-shaped hexamers analogical to the paranuclei of Aβ, Fig 26B(d). Persistent oligomers of the expected MW ~80 kDa and circular appearance (cf. the insert in Fig 26B(d), diameter ~10 nm) are indeed observed during the polymerization in the presence of heme [326]. On the other hand, the available evidence also points to the existence of independent alternative pathways of aggregation and antiparallel β-structure of oligomers [327–331]. The annular assembly of the disk-shaped paranucleus-like hexamers and the subsequent conformational conversion (antiparallel→parallel β-structure) would yield fibrils consisting of two protofilaments which comprise parallel cross-β sheets in the antiparallel alignment. A recently proposed model of α-synuclein protofibril is consistent with this expectation [332], Fig 26B(e); the β arcade fold of each protofilament may render the fibril prone to dissociation and remodelling [333,334].

(viii) Fibrillization of prions via the antiparallel coiled-coil pathway. Fibrillization of the disordered amyloidogenic region of mammalian prions appears to be controlled by the C-terminal well-structured half of the molecule. This part of the molecule is altogether missing in the recombinant polypeptide hPrP23-144 which corresponds to the human PrP Y145Stop variant (the mutation eliminates the entire α-helical region of hPrP), and hPrP23-144 undergoes a spontaneous conversion from the monomeric disordered state to a parallel, in-register, cross-β fibrillar form. The C-terminal I138-F141 segment (red outline in the FP_i plot in Fig 26C(a)), is essential to this conversion but the palindromic M113-A120 segment (blue outline) is not [335–337]. Thus, the anticipated β structure configuration of the amyloidogenic region can be described by the C_{7eq}-C_5-C_5^*-C_5 sequence of 'strands' by analogy to Aβ, tau, and α-synuclein, see the β0-β1-β2-β3 strands in Fig 26C(a). This pattern suggests that hPrP23-144 also forms the antiparallel coiled-coil dimers which assemble via domain swapping into superhelical paranuclei and yield fibrils comprising sheets of parallel cross-β structure.

In the full-length molecule, the same mode of folding can be induced by reduction of pH. The FP_i vs. $\Delta FP_{i-1 \to i+1}$ plots for the native helices of hPrP at high and low pH suggest that the helix α1 is destabilized by reduction of pH while the helices α2 and α3 are becoming stable as the compact structure unravels at low pH. Disruption of the native β-sheet and unfolding of the helix α1 'unmasks' the C_{7eq}-C_5-C_5^*-C_5 sequence, see the $α^i$/A-state in Fig 26C(c). Thus, at low pH, hPrP also forms the antiparallel coiled-coil dimers (B-state) which assemble via domain swapping into superhelical paranuclei like the octamer (β°) shown in Fig 26C(c) [338–348]. The insert shows hPrP aggregates [349] formed, we believe, by tubular stacking of such octamers or similar higher-order oligomers. The stacks of superhelical paranuclei may convert to fibrils comprising two antiparallel sheets of parallel cross-β structure, and the dissociation of such fibrils, combined with conversion of the helices α2 and α3 into extended strands, eventually yields the single-sheet parallel cross-β structure [350,351], Fig 26C(d).

(ix) Fibrillization of prions via the β solenoid pathway. The fibril of GPI-anchorless PrP 27-30 comprises a pair of protofilaments which are formed by stacking four-rung left-handed solenoids that include nearly entire length of the protein [352,353], Fig 26D(a). Thus, the alternative direction of folding of hPrP involves realignment of the native two-stranded β-sheet, and refolding of the three C-terminal helices into a β solenoid while the disulfide bond is intact. The FP_i plot for hPrP (omitting most of the N-terminal domain) suggests that there is indeed the second way to describe the anticipated β structure configuration of the amyloidogenic region, Fig 26D(b). The strands marked by the green outlines in the plot: {β1-β2-β3}

followed by {β4-β5-β6} (i.e. including the palindromic segment) form the sequence of two 'asymmetrically' complementary arrays of alternant incipient strands: C_5-C_{7eq}-C_5^* followed by C_{7eq}-C_5-C_{7eq}. Such a sequence is compatible with the intramolecular parallel rather than anti-parallel assembly of β structure, cf. section **a**.(i) and Figs 2 and 5, and is expected to fold into a two-rung β solenoid. The assignment of the strands marked by the green outlines actually follows the two-rung model [354] proposed as a revision of the original model of the left-handed β helix [355]. The strands β7 through β12 (the 'C_5 strand' segments in the FP_i plot) complete our assignment of the four-rung β solenoid structure which places Cys179 and Cys214 (orange arrows) in register, and all Pro residues (brown arrows) in turns and loops. The assignment implies that the strands β4 and β7 which form the antiparallel β sheet $C_{7eq}\uparrow C_5\downarrow$ in the native structure, do indeed realign into the more stable parallel β sheet $C_{7eq}\uparrow C_5\uparrow$ in the solenoid. The two strands marked by the red outlines, βA and βΩ, have the C_{7eq} and C_5 configurations. The head-to-tail domain swapping between the native conformer and the β solenoid conformer that produces parallel β sheet $\beta A(C_{7eq})\downarrow\beta\Omega(C_5)\downarrow$ (*FP*-directed molecular recognition), may underlie self-propagation of infectious strains of mammalian prions, see Fig 26D(c).

    (x) Fibrillization of HET-s prion forming domains via β solenoid and antiparallel coiled-coil pathways. The HET-s prion forming domain (PDF) from *P. anserina* adopts the two-rung left-handed β solenoid fold in amyloid fibrils [356–358], Fig 26E(a). The FP_i plot at pH 7 confirms that the strands {β1-β1'-β2-β3} and {β4-β4'-β5-β6} of this fold constitute the sequence of two 'asymmetrically' complementary arrays of the alternant incipient strands: C_{7eq}-C_5-C_{7eq}-C_5 followed by C_5-C_{7eq}-C_5-C_{7eq}. As was argued previously, such a sequence is compatible with the intramolecular parallel rather than antiparallel assembly of β structure, and is expected to fold into a two-rung β solenoid as long as such a solenoid fold is supported by the outer side chains' ladders and interior side chains' packing [359–362]. The PFD domain from *F. graminearum* adopts the same fold in spite of limited sequence identity [362] and the FP_i plot reveals similar pattern of the incipient β structure propensity with one exception of the strand β4.

    In contrast, low pH induces formation of a generic amyloid fold of the *P. anserina* domain [363–367] and the M8 variant of this domain forms toxic amyloids comprising antiparallel β structure [368–370]. The FP_i plots in Fig 26E(b) suggest that the wt PFD loses the 'asymmetric' complementarity of the 'strand' sequence, while its variant M8 loses both 'asymmetric' complementarity and the asparagine (Asn) ladders [359–362]. Rather than fold into a solenoid, the protonated or modified HET-s PFD is likely to behave like the C_{7eq}-C_5-C_5^*-C_5 sequence and polymerize via aggregation of the antiparallel dimers [363–370].

Concluding remarks

The qualitative PMO theory of organic structure and reactivity provides simplified models which, inevitably, sacrifice some aspects of reality. The folding potential function, FP_i, that we employed in this study is designed to capture the effect of distribution of charge density on the conformational and H-bonding propensity of the polypeptide backbone; it does not carry any information about the constraints introduced e.g. by the hydrophilic residues or by proline, and obviously does not take into account any side chain-side chain interactions. Indeed, on several occasions in the preceding discussion, it was clearly useful to complement the FP_i plots by showing distribution of the ionized side chains and potential 'steric zipper' segments or the presence of proline and cysteine residues in the sequence. However, simplification may also underlie the strength of the PMO theory models. And so, surprisingly, the FP_i and FP_i vs.$\Delta FP_{i-1\rightarrow i+1}$ plots are found to identify basic features of secondary and tertiary structure, establish sequence correlates of metamorphic and pH-driven equilibria, relate binding affinities and

folding rate constants to secondary-structure preferences, *and manifest common backbone polarization patterns in amyloidogenic regions* of Alzheimer's Aβ and tau, Parkinson's α-synuclein and TSEs' prions. Thus, by stripping away the side chain-side chain interactions, we discover that the main chain H-bonding and hyperconjugation, described in terms of two-electron stabilizing interactions of the filled and vacant molecular orbitals [7–15], make significant and sometimes decisive contributions to the stability of both native and transition states. The new vocabulary and grammar of encoding the 3D structure of proteins which emerge amidst these findings readily define hitherto elusive pathways of misfolding and aggregation associated with brain proteinopathies. One cannot account, it appears, for a wide range of folding phenomena without taking into consideration the stereoelectronic effects such as the generalized anomeric effect and the homohyperconjugation of peptide linkages [16–22,28–33,51–58]. This conclusion is in line with the arguments of backbone-based theory of protein folding [2], and particularly concerns conformational behaviour of highly pleiomorphic proteins associated with neurodegenerative diseases.

Supporting information

S1 Appendix. Folding template and three-dimensional structure of soluble globular proteins.
(PDF)

S2 Appendix. Electronic configuration of the peptide amide bonds and conformational and H-bonding propensity of the polypeptide backbone.
(PDF)

S3 Appendix. Electronic configuration of the polypeptide backbone and canonical peptide recognition by the PDZ domains.
(PDF)

S4 Appendix. Folding template and tertiary structure of soluble proteins: The congeners of *Borrelia* spirochete antigen OspA.
(PDF)

S5 Appendix. Electronic configuration of the polypeptide backbone and molecular recognition in formation of β structure: Folding of inteins.
(PDF)

S1 Table. NMR shielding tensors of C^α atoms of Xaa $\sigma(C^\alpha)^{Xaa}$ (ppm, GIAO//BLYP3/ D95**).
(PDF)

S1 Dataset. Cartesian atomic coordinates and total energies: Helix & Hairpin models (141 entries), secondary structure models 1–5 (5 entries), β structure models 6–8 (8 entries), and Trp-Cage mini-protein model 9 (2 entries).
(PDF)

Acknowledgments

the Department of Chemistry, Yale University, for the support and access to the departmental computing facilities.

Author Contributions

Conceptualization: Andrzej Stanisław Cieplak.

Data curation: Andrzej Stanisław Cieplak.

Formal analysis: Andrzej Stanisław Cieplak.

Investigation: Andrzej Stanisław Cieplak.

Methodology: Andrzej Stanisław Cieplak.

Project administration: Andrzej Stanisław Cieplak.

Writing – original draft: Andrzej Stanisław Cieplak.

Writing – review & editing: Andrzej Stanisław Cieplak.

References

1. Baldwin RL, Rose GD. Is protein folding hierarchic? I. Local structure and peptide folding. Trends Biochem Sci. 1999; 24: 26–33. PMID: 10087919

2. Rose GD, Fleming PJ, Banavar JR, Maritan A. A backbone-based theory of protein folding. Proc Natl Acad Sci USA. 2006; 103: 16623–16633. https://doi.org/10.1073/pnas.0606843103 PMID: 17075053

3. Brogan APS, Siligardi G, Hussain R, Perriman AW, Mann S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. Chem Sci. 2012; 3: 1839–1846.

4. Brogan APS, Sharma KP, Perriman AW, Mann S. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150˚C. Nat Commun. 2014; 5: 5058. https://doi.org/10.1038/ncomms6058 PMID: 25284507

5. Shelimov KB, Jarrold MF. Conformations, unfolding, and refolding of apomyoglobin in vacuum: an activation barrier for gas-phase protein folding. J Am Chem Soc. 1997; 119: 2987–2994;

6. Jarrold MF. Unfolding, refolding, and hydration of proteins in the gas phase. Acc Chem Res. 1999; 32: 360–367.

7. Anslyn EV, Dougherty DA. Modern Physical Organic Chemistry. Sausalito: University Science Books; 2006.

8. Epiotis ND, Cherry WR, Shaik S, Yates RL, Bernardi F. Structural theory in organic chemistry. Top Curr Chem. 1977; 70.

9. Cieplak AS. Stereochemistry of nucleophilic additions to cyclohexanone. The importance of two-electron stabilizing interactions. J Am Chem Soc. 1981; 103: 4540–4552.

10. Spivey AC. Introduction to stereoelectronics. www.imperial.ac.uk/imperial-college/research-centres. . ./lecture10607.pdf

11. Pross A, Radom L. A theoretical approach to substituent interactions in substituted benzenes. In: Taft RW, editor. Progress in Physical Organic Chemistry Vol 13. New York: John Wiley & Sons; 1981. Chapter 1.

12. Weinhold F, Landis CR. Valency and Bonding. A Natural Bond Orbital Donor-Acceptor Perspective. Cambridge: Cambridge University Press; 2006.

13. Fleming I. Molecular Orbitals and Organic Chemical Reactions. Reference Edition. Chichester: Wiley; 2010.

14. Gleiter R, Haberhauer G. Aromaticity and Other Conjugation Effects. Weinheim: Wiley-VCH; 2012.

15. Wu JI, Wang C, McKee WC, Schleyer PvR, Wu W, Mo Y. On the large σ-hyperconjugation in alkanes and alkenes. J Mol Model. 2014; 20: 2228–2236. https://doi.org/10.1007/s00894-014-2228-2 PMID: 24912591

16. Cieplak AS. Solid-state conformations of linear oligopeptides and aliphatic amides: a model of chiral perturbation of a conjugated system. J Am Chem Soc. 1985; 107: 271–273.

17. DeRider ML, Wilkens SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT, Markley JL. Collagen stability: insights from NMR spectroscopy and hybrid density functional computational investigations of

the effect of electronegative substituents on prolyl ring conformations. J Am Chem Soc. 2002; 124: 2497–2505. PMID: 11890798

18. Hinderaker MP, Raines RT, An electronic effect on protein structure. Protein Sci. 2003; 12: 1188–1194. https://doi.org/10.1110/ps.0241903 PMID: 12761389

19. Bartlett GJ, Choudhury A, Raines RT, Woolfson RDN, n−π* interactions in proteins. Nat Chem Biol. 2010; 6: 615–620. https://doi.org/10.1038/nchembio.406 PMID: 20622857

20. Fufezan C. The role of Bürgi-Dunitz interactions in the structural stability of proteins. Proteins. 2010; 78: 2831–2838. https://doi.org/10.1002/prot.22800 PMID: 20635415

21. Wathen B, Pratt DA, Jia Z. Hyperconjugation contributes to the bimodal distribution of glycine conformations observed in protein three-dimensional structures. ChemBioChem. 2011; 12: 1674–1677. https://doi.org/10.1002/cbic.201100156 PMID: 21671332

22. Adhikary R, Zimmermann J, Liu J, Forrest RP, Janicki TD, Dawson PE, Corcelli SA, Romesberg FE. Evidence of an unusual N-H⋯N hydrogen bond in proteins. J Am Chem Soc. 2014; 136: 13474–13477. https://doi.org/10.1021/ja503107h PMID: 25226114

23. Hoffman R. Interactions of orbitals through space and through bonds. Acc Chem Res. 1971; 4: 1–9.

24. Hoffman R. Qualitative thinking in the age of modern computational chemistry. J Mol Struct (Theochem). 1998; 424: 1–6.

25. Dewar MJS, Dougherty RC. The PMO Theory of Organic Chemistry. Plenum Publishing Corporation; 1975.

26. Pauling L, Corey RB. Stable configurations of polypeptide chains. Proc Roy Soc London Series B. 1953; 141: 21–33.

27. Ramakrishnan C, Ramachandran GN. Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. Biophys J. 1965; 5: 909–933. https://doi.org/10.1016/S0006-3495(65)86759-5 PMID: 5884016

28. Cieplak AS. From crystal structures of oligopeptides to protein folding. The importance of peptide bond-side chain hyperconjugation. In: Greenberg A, Breneman C, Liebman JF, editors. The Amide Linkage: Structural Significance in Chemistry, Biochemistry and Materials Science. New York: Wiley; 1999. Chapter 17.

29. Cieplak AS. Inductive and resonance effects of substituents on π-face selection. Chem Rev. 1999; 99: 1265–1336. PMID: 11749447

30. Cieplak AS, Atasoğlu O. Are stereoelectronic effects important in conformational transitions of polypeptide chains? Dependence of nonplanarity of peptide bonds on protein fold in the crystal structures at atomic resolution. IV European Symposium of the Protein Society Paris, 2001. Protein Sci. 2001; 10 Suppl. 1: 111.

31. Cieplak AS, Demir Ö. Extended hyperconjugation of Ala* amino acid side chains: effect on structure of 3-substituted 5,6-diaza-1-bicyclo(2.1.1)hexyl, 5,8-disila-1-bicyclo(2.2.2)octyl and 1bicyclo(2.2.2)octyl cations. IV European Symposium of the Protein Society Paris, 2001. Protein Sci. 2001; 10 Suppl. 1: 146.

32. Cieplak AS, Regression and principal component analyses of internal coordinates for the carboxamides, oligopeptides and lactams. Struct Chem. 1994; 5: 85–98.

33. Cieplak AS, Sürmeli NB, Single-site mutation and secondary structure stability: an isodesmic reaction approach. The case of unnatural amino acid mutagenesis Ala→Lac. J Org Chem. 2004; 69: 3250–3261. https://doi.org/10.1021/jo0358372 PMID: 15132529

34. Banavar JR, Maritan A. Physics of proteins. Annu Rev Biophys Biomol Struct. 2007; 36: 261–280. https://doi.org/10.1146/annurev.biophys.36.040306.132808 PMID: 17477839

35. Haddon RC, Perturbational molecular orbital (PMO) theory of homoaromaticity. J Am Chem Soc. 1975; 97: 3608–3615.

36. Sippl MJ. Fold space unlimited. Curr Opin Struc Biol. 2009; 19: 312–320.

37. Murzin AG. Metamorphic proteins. Science. 2008; 320: 1725–1726. https://doi.org/10.1126/science.1158868 PMID: 18583598

38. Baranczak A, Kelly JW, A current pharmacologic agent versus the promise of the next generation therapeutics to ameliorate protein misfolding and/or aggregation diseases. Curr Opin Chem Biol. 2016; 32: 10–21. https://doi.org/10.1016/j.cbpa.2016.01.009 PMID: 26859714

39. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing protein intrinsic disorder. Chem Rev. 2014; 114: 6561–6588. https://doi.org/10.1021/cr400514h PMID: 24739139

40. Qing G, Zhao S, Xiong Y, Jiang F, Liu Y, Chen H, et al. Chiral effect at protein/graphene interface: A bioinspired perspective to understand amyloid formation. J Am Chem Soc. 2014; 136: 10736–10742. https://doi.org/10.1021/ja5049626 PMID: 25011035

41. Gao G, Zhang M, Lu P, Guo G, Wang D, Sun T. Chirality-assisted ring-like aggregation of Aβ(1–40) at liquid-solid interfaces: A stereoselective two-step assembly process. Angew Chem Int Ed. 2015; 54: 2245–2250.

42. Kowalewski T, Holtzman DM. *In situ* atomic force microscopy study of Alzheimer's β-amyloid peptide on different substrates: New insights into mechanism of β-sheet formation. Proc Natl Acad Sci USA. 1999; 96: 3688–3693. PMID: 10097098

43. Meisl G, Yang X, Hellstrand E, Frohm B, Kirkegaard JB, Cohen SAI et al. Differences in nucleation behaviour underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. Proc Natl Acad Sci USA. 2014; 111: 9384–9389. https://doi.org/10.1073/pnas.1401564111 PMID: 24938782

44. Albright TA, Burdett J, Whangbo MH. Orbital Interactions in Chemistry. 2nd ed. New York: John Wiley & Sons; 2013.

45. Ho BK, Thomas A, Brasseur R. Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatic, and H-bonding in the α-helix. Protein Sci. 2003; 12: 2508–2522. https://doi.org/10.1110/ps.03235203 PMID: 14573863

46. Mo Y, Gao J. Theoretical analysis of the rotational barrier of ethane. Acc Chem Res. 2007; 40: 113–119. https://doi.org/10.1021/ar068073w PMID: 17309192

47. Onsager L. Electric moments of molecules in liquids. J Am Chem Soc. 1936; 58: 1486–1493.

48. Debye P, Hückel E. Zur Theorie der Elektrolyte. I. Gefrierpunktsernidrigung und verwandte Erscheinungen. Phys Zeit. 1923; 24: 185–206.

49. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 98 Revision A.3. Pittsburgh: Gaussian Inc; 1998.

50. Neidigh JW, Fesinmeyer RM, Andersen NH. Designing a 20-residue protein. Nat Struct Biol. 2002; 9: 425–430. https://doi.org/10.1038/nsb798 PMID: 11979279

51. Fleming I. op. cit. pp 96–112. Cf. structures 2.151 and 2.152.

52. Shamovsky IL, Ross GM, Riopelle RJ. Theoretical studies on the origin of β-sheet twisting. J Phys Chem B. 2000; 104: 11296–11307.

53. Fischer FR, Wood PA, Allen FH, Diederich F. Orthogonal dipolar interactions between amide carbonyl groups. Proc Natl Acad Sci USA. 2008; 105: 17290–17294. https://doi.org/10.1073/pnas.0806129105 PMID: 18981424

54. Chakrabarti P, Pal D. An electrophile-nucleophile interaction in metalloprotein structures. Protein Sci. 1997; 6: 851–859. https://doi.org/10.1002/pro.5560060412 PMID: 9098895

55. Cieplak AS. Organic addition and elimination reactions: transformation paths of carbonyl derivatives. In: Dunitz J, Bürgi HB, editors. Structure Correlation Vol 1. Weinheim: VCH; 1994. Chapter 6.

56. Worley B, Richard G, Harbison GS, Powers R. ^{13}C NMR reveals no evidence of $n{-}\pi^*$ interactions in proteins. PLoS ONE. 2012; 7: e42075. https://doi.org/10.1371/journal.pone.0042075 PMID: 22876300

57. Kamer KJ, Choudhary A, Raines RT. Intimate interactions with carbonyl groups: dipole-dipole or $n{\to}\pi^*$? J Org Chem. 2013; 78: 2099–2103. https://doi.org/10.1021/jo302265k PMID: 23163960

58. Newberry RW, Bartlett GJ, VanVeller B, Woolfson DN, Raines RT, Signatures of $n{\to}\pi^*$ interactions in proteins. Protein Sci. 2014; 23; 284–288. https://doi.org/10.1002/pro.2413 PMID: 24375625

59. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. Prog Biophys Mol Biol. 1984; 44: 97–179. PMID: 6385134

60. Wiberg KB. Origin of the amide rotational barrier. In: Greenberg A, Breneman C, Liebman JF, editors. The Amide Linkage: Structural Significance in Chemistry, Biochemistry and Materials Science. New York: Wiley; 2003. Chapter 2.

61. Lario PI, Vrielink A. Atomic resolution density maps reveal secondary structure dependent differences in electronic distribution. J Am Chem Soc. 2003; 125: 12787–12794. https://doi.org/10.1021/ja0289954 PMID: 14558826

62. Milner-White EJ. The partial charge of the nitrogen atom in peptide bonds. Protein Sci. 1997; 6: 2477–2482. https://doi.org/10.1002/pro.5560061125 PMID: 9385654

63. Glendening ED, Hrabal JA II. Resonance in formamide and its chalcogen replacement analogues: a Natural Population Analysis/Natural Resonance Theory viewpoint. J Am Chem Soc. 1997; 119: 12940–12946.

64. Klotz IM, Franzen JS. Hydrogen bonds between model peptide groups in solution. J Am Chem Soc. 1962; 84: 3461–3466.

65. Ben-Tal N, Sitkoff D, Topol IA, Yang AS, Burt SK, Honig B. Free energy of amide hydrogen bond formation in vacuum, in water, and in liquid alkane solution. J Phys Chem B. 1997; 101: 450–457.

66. Derewenda ZS, Lee L, Derewenda U. The occurrence of C-H···O hydrogen bonds in proteins. J Mol Biol. 1995; 252: 248–262. PMID: 7674305

67. Fabiola GF, Krishnaswamy S, Nagarajan V, Pattabhi V. C-H···O hydrogen bonds in β-sheets. Acta Crystallogr D. 1997; 53: 316–320. https://doi.org/10.1107/S0907444997000383 PMID: 15299935

68. Penel S, Morrison GM, Dobson PD, Mortshire-Smith RJ, Doig AJ. Length preferences and periodicity in β-strands. Antiparallel edge β-sheets are more likely to finish in non-hydrogen bonded rings. Protein Eng. 2003; 16: 957–961. https://doi.org/10.1093/protein/gzg147 PMID: 14983075

69. Sridhar S, Guruprasad K. Can natural proteins designed with 'inverted' peptide sequences adopt native-like protein folds? PLoS ONE. 2014; 9: e107647. https://doi.org/10.1371/journal.pone.0107647 PMID: 25210740

70. Banerjii-Basu S, Baxevanis AD. Molecular evolution of the homeodomain family of transcription factors. Nucl Acid Res. 2001; 29: 3258–3269.

71. Griffiths-Jones SR, Searle MS. Structure, folding, and energetics of cooperative interactions between the β-strands of a *de novo* designed three-stranded antiparallel β-sheet peptide. J Am Chem Soc. 2000; 122: 8350–8356.

72. de Alba E, Santoro J, Rico M, Jimenez MA. De novo design of a monomeric three-stranded antiparallel β-sheet. Protein Sci. 1999; 8: 854–865. https://doi.org/10.1110/ps.8.4.854 PMID: 10211831

73. Lopez de la Paz M, Lacroix E, Ramirez-Alvarado M, Serrano L. Computer-aided design of β-sheet peptides. J Mol Biol. 2001; 312: 229–246. https://doi.org/10.1006/jmbi.2001.4918 PMID: 11545599

74. Fisk JD, Schmitt MA, Gellman SH. Thermodynamic analysis of autonomous parallel β-sheet formation in water. J Am Chem Soc. 2006; 128: 7148–7149. https://doi.org/10.1021/ja060942p PMID: 16734453

75. Kung VM, Cornilescu G, Gellman SH. Impact of strand number on parallel β-sheet stability. Angew Chem Int Ed. 2015; 54: 14336–14339.

76. Wiberg KB, Rablen PR, Rush DJ, Keith TA. Amides. 3. Experimental and theoretical studies of the effect of the medium on the rotational barriers for N,N-dimethylformamide and N,N-dimethylacetamide. J Am Chem Soc. 1995; 117: 4261–4270.

77. Kukic P, Farrell D, McIntosh LP, Garcia-Moreno EB, Steen Jensen K, Toleikis Z, et al. Protein dielectric constants determined from NMR chemical shift perturbation. J Am Chem Soc. 2013; 135: 16968–16976. https://doi.org/10.1021/ja406995j PMID: 24124752

78. Cuervo A, Dans PD, Carrascosa JL, Orozco M, Gomila G, Fumagalli L. Direct measurement of the dielectric polarization properties of DNA. Proc Natl Acad Sci USA. 2014; 111: E3624–E3630. https://doi.org/10.1073/pnas.1405702111 PMID: 25136104

79. Rajan R, Awasthi SK, Bhattacharjya S, Balaram P. "Teflon-coated peptides": Hexafluoroacetonetrihydrate as a structure stabilizer for peptides. Biopolymers. 1997; 42: 125–128. https://doi.org/10.1002/(SICI)1097-0282(199708)42:2<125::AID-BIP1>3.0.CO;2-P PMID: 9234991

80. Siglioccolo A, Paiardini A, Piscitelli M, Pascarella S. Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. BMC Struct Biol. 2011; 11: 50. https://doi.org/10.1186/1472-6807-11-50 PMID: 22192175

81. Elgabarty H, Khaliullin RZ, Kühne TD. Covalency of hydrogen bonds in liquid water can be probed by proton nuclear magnetic resonance experiments. Nat Commun. 2015; 6: 8318. https://doi.org/10.1038/ncomms9318 PMID: 26370179

82. Bockris JO'M, Reddy AKN. Modern Electrochemistry Vol 1: Ionics. New York: Plenum Press; 1998.

83. Ghosh JC. The abnormality of strong electrolytes. Part 1. Electrical conductivity of aqueous salt solutions. J Chem Soc Trans. 1918; 113: 449–458.

84. Kendall J. The abnormality of strong electrolytes and the ionization theory of Ghosh. J Am Chem Soc. 1922; 44: 717–738.

85. Steinmetz MO, Stock A, Schulthess T, Landwehr R, Lustig A, Faix J, et al. A distinct 14 residue site triggers coiled-coil formation in cortexillin I. EMBO J. 1998; 17: 1883–1891. https://doi.org/10.1093/emboj/17.7.1883 PMID: 9524112

86. Spitzer J, Poolman B. The role of biomacromolecular crowding, ionic strength, and physicochemical gradients in the complexities of life's emergence. Microbiol Mol Biol Rev. 2009; 73: 371–388. https://doi.org/10.1128/MMBR.00010-09 PMID: 19487732

87. Kragelund B, Højrup P, Skou Jensen M, Karlskov Schjerling C, Juue E, Knudsen J, et al. Fast and one-step folding of closely and distantly related homologous proteins of a four-helix bundle family. J Mol Biol. 1996; 256: 187–200. PMID: 8609609

88. Burton RE, Huang GS, Daugherty MA, Fullbright PW, Oas TG. Microsecond protein folding through a compact transition state. J Mol Biol. 1996; 263: 311–322; https://doi.org/10.1006/jmbi.1996.0577 PMID: 8913309

89. Ferguson N, Capaldi AP, James R, Kleanthous C, Radford SE. Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. J Mol Biol. 1999; 286: 1597–1608. https://doi.org/10.1006/jmbi.1998.2548 PMID: 10064717

90. Mayor U, Johnson CM, Daggett V, Fersht AR. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. Proc Natl Acad Sci USA. 2000; 97: 13518–13522. https://doi.org/10.1073/pnas.250473497 PMID: 11087839

91. Myers JK, Oas TG. Preorganized secondary structure as an important determinant of fast protein folding. Nat Struct Biol. 2001; 8: 552–558. https://doi.org/10.1038/88626 PMID: 11373626

92. Ferguson N, Li W, Capaldi AP, Kleanthous C, Radford SE. Using chimeric immunity proteins to explore the energy landscape for α-helical protein folding. J Mol Biol. 2001; 307: 393–405. https://doi.org/10.1006/jmbi.2000.4492 PMID: 11243827

93. Kubelka J, Eaton WA, Hofrichter J. Experimental test of villin subdomain folding simulations. J Mol Biol. 2003; 329: 625–630. PMID: 12787664

94. Zhu Y, Alonso DOV, Maki K, Huang CY, Lahr SJ, Daggett V, et al. Ultrafast folding of α₃D: A de novo designed three-helix bundle protein. Proc Natl Acad Sci USA. 2003; 100: 15486–15491. https://doi.org/10.1073/pnas.2136623100 PMID: 14671331

95. Gianni S, Guydosh NR, Khan F, Caldas TD, Mayor U, White GWN, et al. Unifying features in protein folding mechanism. Proc Natl Acad Sci USA. 2003; 100: 13286–13291. https://doi.org/10.1073/pnas.1835776100 PMID: 14595026

96. Zhu YJ, Fu X, Wang T, Tamura A, Takada S, Saven JG, et al. Guiding the search for a protein's maximum rate of folding. Chem Phys. 2004; 307: 99–109.

97. Dimitriadis G, Drysdale A, Myers JK, Arora P, Radford SE, Oas TG, et al. Microsecond folding dynamics of the F13WG29A mutant of the B domain of staphylococcal protein A by laser-induced temperature jump. Proc Natl Acad Sci USA. 2004; 101: 3809–3814. https://doi.org/10.1073/pnas.0306433101 PMID: 15007169

98. Jemth P, Day R, Gianni S, Khan F, Allen M, Daggett V, et al. The structure of the major transition state for folding of an FF domain from experiment and simulation. J Mol Biol. 2005; 350: 363–378. https://doi.org/10.1016/j.jmb.2005.04.067 PMID: 15935381

99. Giri R, Morrone A, Travaglini-Allocatelli C, Jemth P, Brunori M, Gianni S, Folding pathways of proteins with increasing degree of sequence identities but different structure and function. Proc Natl Acad Sci USA. 2012; 109: 17772–17776. https://doi.org/10.1073/pnas.1201794109 PMID: 22652570

100. Dodson CA, Arbely E. Protein folding of the SAP domain, a naturally occurring two-helix bundle. FEBS Lett. 2015; 589: 1740–1747. https://doi.org/10.1016/j.febslet.2015.06.002 PMID: 26073259

101. Pace CN, Scholtz JM. A helix propensity scale based on experimental studies of peptides and proteins. Biophys J. 1998; 75: 422–427. PMID: 9649402

102. Lyu P, Sherman JC, Chen A, Kallenbach NR. α-Helix stabilization by natural and unnatural amino acids with alkyl side chains. Proc Natl Acad Sci USA. 1991; 88: 5317–5320. PMID: 2052608

103. Padmanabhan S, Baldwin RL. Straight-chain non-polar amino acids are good helix formers in water. J Mol Biol. 1991; 219: 135–137. PMID: 2038048

104. Chakrabartty A, Kortemme T, Baldwin RL. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. Protein Sci. 1994; 3: 843–852. https://doi.org/10.1002/pro.5560030514 PMID: 8061613

105. Padmanabhan S, York EJ, Stewart JM, Baldwin RL. Helix propensities of basic amino acids increase with the length of the side-chain. J Mol Biol. 1996; 257; 726–734. https://doi.org/10.1006/jmbi.1996.0197 PMID: 8648636

106. Andrew CD, Warwicker J, Jones GR, Doig AJ. Effect of phosphorylation on α-helix stability as a function of position. Biochemistry. 2002; 41: 1897–1905. PMID: 11827536

107. Chiu HP, Suzuki Y, Gullickson D, Ahmad R, Kokona B, Fairman R, et al. Helix propensity of highly fluorinated amino acids. J Am Chem Soc. 2006; 128: 15556–15557. https://doi.org/10.1021/ja0640445 PMID: 17147342

108. Cheng RP, Weng YJ, Wang WR, Koyack MJ, Suzuki Y, Wu CH, et al. Helix formation and capping energetics of arginine analogs with varying side chain lengths. Amino Acids. 2012; 43: 195–206. https://doi.org/10.1007/s00726-011-1064-2 PMID: 21922267

109. Lawrence JR, Johnson WC. Lifson-Roig nucleation for α-helices in trifluoroethanol: context has a strong effect on the helical propensity of amino acids. Biophys Chem. 2002; 101–102: 375–385. PMID: 12488015

110. Kim CA, Berg JM. Thermodynamic β-sheet propensities measured using a zinc-finger host peptide. Nature. 1993; 362: 267–270. https://doi.org/10.1038/362267a0 PMID: 8459852

111. Minor DL Jr, Kim PS. Context is a major determinant of β-sheet propensity. Nature. 1994; 371: 264–267. https://doi.org/10.1038/371264a0 PMID: 8078589

112. Minor DL Jr, Kim PS. Measurement of the β-sheet-forming propensities of amino acids. Nature. 1994; 367: 660–663. https://doi.org/10.1038/367660a0 PMID: 8107853

113. Smith CK, Withka JM, Regan L. A thermodynamic scale for the β-sheet forming tendencies of the amino acids. Biochemistry. 1994; 33: 5510–5517. PMID: 8180173

114. Chiu HP, Kokona B, Fairman R, Cheng RP. Effect of highly fluorinated amino acids on protein stability at a solvent-exposed position on an internal strand of protein G B1 domain. J Am Chem Soc. 2009; 131: 13192–13193. https://doi.org/10.1021/ja903631h PMID: 19711980

115. Natesh R, Manikandan K, Bhanumoorthy P, Viswamitra MA, Ramakumar S. Thermostable xylanase from *Thermoascus auranticus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anistropically to small-molecule accuracy. Acta Crystallogr D. 2003; 59: 105–117. PMID: 12499546

116. Deechongkit S, Nguyen H, Powers ET, Dawson PE, Gruebele M, Kelly JW. Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. Nature. 2004; 430: 101–105. https://doi.org/10.1038/nature02611 PMID: 15229605

117. Deechongkit S, Dawson PE, Kelly JW. Toward assessing the position-dependent contributions of backbone hydrogen bonding to β-sheet folding thermodynamics employing amide-to-ester perturbations. J Am Chem Soc. 2004; 126: 16762–16771. https://doi.org/10.1021/ja045934s PMID: 15612714

118. Beerten J, Van Durme J, Gallardo R, Capriotti E, Serpell L, Rousseau F, et al. WALTZ-DB: a benchmark database of amyloidogenic hexapeptides. Bioinformatics. 2015; 31: 1698–1700. https://doi.org/10.1093/bioinformatics/btv027 PMID: 25600945

119. Eildal JNN, Hultqvist G, Balle T, Stuhr-Hansen N, Padrah S, Gianni S, et al. Probing the role of backbone hydrogen bonds in protein-peptide interactions by amide-to-ester mutations. J Am Chem Soc. 2013; 135: 12998–13007. https://doi.org/10.1021/ja402875h PMID: 23705582

120. Stiffler MA, Chen JR, Grantcheva VP, Lei Y, Fuchs D, Allen JE, et al. PDZ domain binding is optimized across the mouse proteome. Science. 2007; 317: 364–369. https://doi.org/10.1126/science.1144592 PMID: 17641200

121. Wiedemann U, Boisguerin P, Leben R, Leitner D, Krause G, Moelling K, et al. Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides. J Mol Biol. 2004; 343: 703–718. https://doi.org/10.1016/j.jmb.2004.08.064 PMID: 15465056

122. Shepherd TR, Hard RL, Murray AM, Pei D, Fuentes EJ. Distinct ligand specificity of the Tiam1 and Tiam2 PDZ domains. Biochemistry. 2011; 50: 1296–1308. https://doi.org/10.1021/bi1013613 PMID: 21192692

123. Saro D, Li T, Rupasinghe C, Paredes A, Caspers N, Spaller MR. A thermodynamic ligand binding study of the third PDZ domain (PDZ3) from the mammalian neuronal protein PSD-95. Biochemistry. 2007; 46: 6340–6352. https://doi.org/10.1021/bi062088k PMID: 17474715

124. Pace CN, Scholtz JM, Grimsley GR. Forces stabilizing proteins. FEBS Lett. 2014; 588: 2177–2184. https://doi.org/10.1016/j.febslet.2014.05.006 PMID: 24846139

125. Schortle D, Stites WE, Meeker AK. Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. Biochemistry. 1990; 29: 8033–8041. PMID: 2261461

126. Pace CN, Fu H, Lee Fryar K, Landua J, Trevino SR, Shirley BA, et al. Contributions of hydrophobic interactions to protein stability. J Mol Biol. 2011; 408: 514–528. https://doi.org/10.1016/j.jmb.2011.02.053 PMID: 21377472

127. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, et al. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science. 1992; 255: 178–183. PMID: 1553543

128. Xu J, Baase WA, Baldwin E, Matthews BW. The response of T4 lysozyme to large-to-small substitutions within its core and its relation to the hydrophobic effect. Protein Sci. 1998; 7: 158–177. https://doi.org/10.1002/pro.5560070117 PMID: 9514271

129. Hamill SJ, Meekhof AE, Clarke J. The effect of boundary selection on the stability and folding of the third fibronectin type III domain from human tenascin. Biochemistry. 1998; 37: 8071–8079. https://doi.org/10.1021/bi9801659 PMID: 9609701

130. Cota E, Hamill SJ, Fowler SB, Clarke J. Two proteins with the same structure respond very differently to mutation: the role of plasticity in protein stability. J Mol Biol. 2000; 302: 713–725. https://doi.org/10.1006/jmbi.2000.4053 PMID: 10986129

131. Main ERG, Fulton KF, Jackson SE. Context-dependent nature of destabilizing mutations on the stability of FKPB12. Biochemistry. 1998; 37: 6145–6153. https://doi.org/10.1021/bi973111s PMID: 9558354

132. Li SC, Deber CM. A measure of helical propensity for amino acids in membrane environments. Nat Struc Biol. 1994; 1: 368–373.

133. Jarrold MF. Helices and sheets *in vacuo*. Phys Chem Chem Phys. 2007; 9: 1659–1671. https://doi.org/10.1039/b612615d PMID: 17396176

134. Lipfert J, Doniach S, Das R, Herschlag D. Understanding nucleic acid-ion interactions. Annu Rev Biochem. 2014; 83: 813–841. https://doi.org/10.1146/annurev-biochem-060409-092720 PMID: 24606136

135. Guéroult M, Boittin O, Mauffret O, Etchebest C, Hartmann B. Mg^{2+} in the major groove modulates B-DNA structure and dynamics. PLoS ONE. 2012; 7: e41704. https://doi.org/10.1371/journal.pone.0041704 PMID: 22844516

136. Sheridan RP, Lee RH, Peters N, Allen LC. Hydrogen-bonded cooperativity in protein secondary structure. Biopolymers. 1979; 18: 2451–2458.

137. Harder E, Anisimov VM, Whitfield T, MacKerrel AD Jr, Roux B. Understanding the dielectric properties of liquid amides from a polarizable force field. J Phys Chem B. 2008; 112: 3509–3521. https://doi.org/10.1021/jp709729d PMID: 18302362

138. Zandomeneghi G, Krebs MRH, McCammon MG, Fändrich M. FTIR reveals differences between native β-sheet proteins and amyloid fibrils. Protein Sci. 2004; 13: 3314–3321. https://doi.org/10.1110/ps.041024904 PMID: 15537750

139. Shivu B, Seshadri S, Li J, Oberg KA, Uversky VN, Fink AL. Distinct β-sheet structure in protein aggregates determined by ATR-FTIR spectroscopy. Biochemistry. 2013; 52: 5176–5183. https://doi.org/10.1021/bi400625v PMID: 23837615

140. Mantovani A, Bonecchi R, Locati M. Tuning inflammation and immunity by chemokine sequestration: decoys and more. Nat Rev Immun. 2006; 6: 907–918.

141. Exner V, Aichinger E, Shu H, Wildhaber T, Alfarano P, Caflisch A, et al. The chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 is essential for H3K27me3 binding and function in Arabidopsis development. PLoS ONE. 2009; 4: e5335. https://doi.org/10.1371/journal.pone.0005335 PMID: 19399177

142. Krishna N, Guruprasad K. Certain heptapeptide and large sequences representing an entire helix, strand or coil conformation in proteins are associated as chameleon sequences. Int J Biol Macromol. 2011; 49: 218–222. https://doi.org/10.1016/j.ijbiomac.2011.04.017 PMID: 21569793

143. Anderson WJ, Van Dorn LO, Ingram WM, Cordes MHJ. Evolutionary bridges to new protein folds: design of C-terminal protein chameleon sequences. Protein Eng. 2011; 24: 765–771.

144. Kuloglu ES, McCaslin DR, Kitabwalla M, Pauza CD, Markley JL, Volkman BF. Monomeric solution structure of the prototypical 'C' chemokine lymphotactin. Biochemistry. 2001; 40: 12486–12496. PMID: 11601972

145. Tuinstra RL, Peterson FC, Kutlesa S, Elgin ES, Kron MA, Volkman BF. Interconversion between two unrelated protein folds in the lymphotactin native state. Proc Natl Acad Sci USA. 2008; 105: 5057–5062. https://doi.org/10.1073/pnas.0709518105 PMID: 18364395

146. He Y, Chen Y, Alexander PA, Bryan PN, Orban J. Mutational tipping points for switching protein folds and functions. Structure. 2012; 20: 283–291. https://doi.org/10.1016/j.str.2011.11.018 PMID: 22325777

147. Alexander PA, He Y, Chen Y, Orban J, Bryan PN. A minimal sequence code for switching protein structure and function. Proc Natl Acad Sci USA. 2009; 106: 21149–21154. https://doi.org/10.1073/pnas.0906408106 PMID: 19923431

148. Luo X, Tang Z, Xia G, Wassmann K, Matsumoto T, Rizo J, et al. The Mad2 spindle checkpoint protein has two distinct natively folded states. Nat Struc Mol Biol. 2004; 11: 338–345.

149. Hsu WL, Oldfield CJ, Xue B, Meng J, Huang F, Romero P, et al. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. Protein Sci. 2013; 22: 258–273. https://doi.org/10.1002/pro.2207 PMID: 23233352

150. De Guzman RN, Martinez-Yamout MA, Dyson HJ, Wright PE. Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2. J Biol Chem. 2004; 279: 3042–3049. https://doi.org/10.1074/jbc.M310348200 PMID: 14594809

151. Kingston RL, Hammel DJ, Gay LS, Dahlquist FW, Matthews BW. Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc Natl Acad Sci USA. 2004; 101: 8301–8306. https://doi.org/10.1073/pnas.0402690101 PMID: 15159535

152. Wang Y, Chu X, Longhi S, Roche P, Han W, Wang E, et al. Multiscaled exploration of coupled folding and binding of an intrinsically disordered molecular recognition element in measles virus nucleoprotein. Proc Natl Acad Sci USA. 2013; 110: E3743–E3752. https://doi.org/10.1073/pnas.1308381110 PMID: 24043820

153. Eryilmaz E, Shah NH, Muir TW, Cowburn D. Structural and dynamical features of inteins and implications on protein splicing. J Biol Chem. 2014; 289: 14506–14511. https://doi.org/10.1074/jbc.R113.540302 PMID: 24695731

154. Salemme FR, Structural properties of protein β-sheets. Prog Biophys Mol Biol. 1983; 42: 95–133. PMID: 6359272

155. Stanger HE, Syud FA, Espinosa JF, Giriat I, Muir T, Gellman SH. Length-dependent stability and strand length limits in antiparallel β-sheet secondary structure. Proc Natl Acad Sci USA. 2001; 98: 12015–12020. https://doi.org/10.1073/pnas.211536998 PMID: 11593011

156. Shah NH, Eryilmaz E, Cowburn D, Muir TW. Naturally split inteins assemble through a 'capture and collapse' mechanism. J Am Chem Soc. 2013; 135: 18673–18681. https://doi.org/10.1021/ja4104364 PMID: 24236406

157. Susa AC, Wu C, Bernstein SL, Dupuis NF, Wong H, Raleigh DP, et al. Defining the molecular basis of amyloid inhibitors: human islet amyloid polypeptide-insulin interactions. J Am Chem Soc. 2014; 136: 12912–12919. https://doi.org/10.1021/ja504031d PMID: 25144879

158. Bullough PA, Hughson FM, Skehel JJ, Wiley DC. Structure of influenza hemagglutinin at the pH of membrane fusion. Nature. 1994; 371: 37–43. https://doi.org/10.1038/371037a0 PMID: 8072525

159. Leka O, Vallese F, Pirazzini M, Berto P, Montecucco C, Zanotti G. Diphteria toxin conformational switching at acidic pH. FEBS J. 2014; 281: 2115–2122. https://doi.org/10.1111/febs.12783 PMID: 24628974

160. Pasquato N, Berni R, Folli C, Alfier B, Cendron L, Zanotti G. Acidic pH-induced conformational changes in amyloidogenic mutant transthyretin. J Mol Biol. 2007; 366: 711–719. https://doi.org/10.1016/j.jmb.2006.11.076 PMID: 17196219

161. Palaninathan SK, Mohamedmohaideen NN, Snee WC, Kelly JW, Sacchettini JC. Structural insight into pH-induced conformational change within the native transthyretin tetramer. J Mol Biol. 2008; 382: 1157–1167. https://doi.org/10.1016/j.jmb.2008.07.029 PMID: 18662699

162. Bateman DA, Tycko R, Wickner RB. Experimentally derived structural constraints for amyloid fibrils of wild-type transthyretin. Biophys J. 2011; 101: 2485–2492. https://doi.org/10.1016/j.bpj.2011.10.009 PMID: 22098747

163. Lim KH, Dasari AKR, Hung I, Gan Z, Kelly JW, Wright PE, et al. Solid-state NMR studies reveal native-like β-sheet structures in transthyretin amyloid. Biochemistry. 2016; 55: 5272–5278. https://doi.org/10.1021/acs.biochem.6b00649 PMID: 27589034

164. Lim KH, Dasari AKR, Hung I, Gan Z, Kelly JW, Wemmer DE. Structural changes associated with transthyretin misfolding and amyloid formation revealed by solution and solid-state NMR. Biochemistry. 2016; 55: 1941–1944. https://doi.org/10.1021/acs.biochem.6b00164 PMID: 26998642

165. Serag AA, Altenbach C, Gingery M, Hubbell WL, Yeates TO. Identification of a subunit interface in transthyretin amyloid fibrils: evidence for self-assembly from oligomeric building blocks. Biochemistry. 2001; 40: 9089–9096. PMID: 11478875

166. Laidman J, Forse GJ, Yeates TO. Conformational change and assembly through edge β strands in transthyretin and other amyloid proteins. Acc Chem Res. 2006; 39: 576–583. https://doi.org/10.1021/ar050017s PMID: 16981673

167. Goldsteins G, Persson H, Andersson K, Olofsson A, Dacklin I, Edvinsson Å, et al. Exposure of cryptic epitopes on transthyretin only in amyloid and in amyloidogenic mutants. Proc Natl Acad Sci USA. 1999; 96: 3108–3113. PMID: 10077645

168. Olofsson A, Ippel JH, Wijmenga SS, Lundgren E, Ohman A. Probing solvent accessibility of transthyretin amyloid by solution NMR spectroscopy. J Biol Chem. 2004; 279: 5699–5707. https://doi.org/10.1074/jbc.M310605200 PMID: 14604984

169. Groenning M, Campos RI, Hirschberg D, Hammarström P, Vestergaard B. Considerably unfolded monomers preceed and exchange with dynamically structured amyloid protofibrils. Sci Rep. 2015; 5: 11443. https://doi.org/10.1038/srep11443 PMID: 26108284

170. Pires RH, Karsai A, Saraiva MJ, Damas AM, Kellermayer SZ, Distinct annular oligomers captured along the assembly and disassembly pathways of transthyretin amyloid protofibrils. PLoS ONE. 2012; 7: e44992. https://doi.org/10.1371/journal.pone.0044992 PMID: 22984597

171. Faria TQ, Almeida ZL, Cruz PF, Jesus CSH, Castanheira P, Brito RMM. A look into amyloid formation by transthyretin: aggregation pathway and a novel kinetic model. Phys Chem Chem Phys. 2015; 17: 7255. https://doi.org/10.1039/c4cp04549a PMID: 25694367

172. Pires RH, Saraiva MJ, Damas AM, Kellermayer SZ. Force spectroscopy reveals the presence of structurally modified dimers in transthyretin amyloid annular oligomers. J Mol Recognit. 2017; 30: e2587.

173. Roychaudhuri R, Yang M, Hoshi MM, Teplow DB. Amyloid β-protein assembly and Alzheimer disease. J Biol Chem. 2009; 284: 4749–4753. https://doi.org/10.1074/jbc.R800036200 PMID: 18845536

174. Ma K, Clancy EL, Zhang Y, Ray DG, Wollenberg K, Zagorski MG. Residue-specific pK_a measurements of the β-peptide and mechanism of pH-induced amyloid formation. J Am Chem Soc. 1999; 121: 8698–8706.

175. Zirah S, Kozin SA, Mazur AK, Blond A, Cheminant M, Ségalas-Milazzo I, et al. Structural changes of region 1–16 of the Alzheimer disease amyloid β-peptide upon zinc binding and *in vitro* aging. J Biol Chem. 2006; 281: 2151–2161. https://doi.org/10.1074/jbc.M504454200 PMID: 16301322

176. Yagi-Utsumi M, Kato K, Nishimura K. Membrane-induced dichotomous conformation of amyloid β with the disordered N-terminal segment followed by the stable C-terminal β structure. PLoS ONE. 2016; 11: e0146405. https://doi.org/10.1371/journal.pone.0146405 PMID: 26731546

177. Yagi-Utsumi M, Matsuo K, Yanagisawa K, Gekko K, Kato K. Spectroscopic characterization of intermolecular interaction of amyloid β promoted on GM1 micelles. Int J Alzheimer's Dis. 2011; 11: 925073.

178. Miles LA, Crespi GA, Doughty L, Parker MW. Bapineuzumab captures the N-terminus of the Alzheimer's disease amyloid-beta peptide in a helical conformation. Sci Rep. 2013; 3: 1302. https://doi.org/10.1038/srep01302 PMID: 23416764

179. Sticht H, Bayer P, Willbold D, Dames S, Hilbich C, Beyreuther K, et al. Structure of amyloid Aβ-(1–40)-peptide of Alzheimer's disease. Eur J Biochem. 1995; 233: 293–298. PMID: 7588758

180. Crescenzi O, Tomaselli S, Guerrini R, Salvadori S, D'Ursi AM, Temussi PA, et al. Solution structure of the Alzheimer amyloid β-peptide (1–42) in an apolar microenvironment. Eur J Biochem. 2002; 269: 5642–5648. PMID: 12423364

181. Coles M, Bicknell W, Watson AA, Fairlie DP, Craik DJ. Solution structure of amyloid β-peptide(1–40) in a water micelle environment. Is the membrane-spanning domain where we think it is? Biochemistry. 1998; 37: 11064–11077. https://doi.org/10.1021/bi972979f PMID: 9693002

182. Shao H, Jao S, Ma K, Zagorski MG. Solution structure of micelle-bound amyloid β-(1–40) and β-(1–42) peptides of Alzheimer's disease. J Mol Biol. 1999; 285: 755–773. PMID: 9878442

183. Mikhalyov I, Olofsson A, Gröbner G, Johansson LBÅ. Designed fluorescent probes reveal interactions between amyloid-β(1–40) peptides and GM1 gangliosides in micelles and lipid vesicles. Biophys J. 2010; 99: 1510–1519. https://doi.org/10.1016/j.bpj.2010.06.043 PMID: 20816063

184. Korshavn KJ, Bhunia A, Lim MH, Ramamoorthy A. Amyloid-β adopts a conserved, partially folded structure upon binding to zwitterionic lipid bilayers prior to amyloid formation. Chem Commun. 2016; 52: 882–885.

185. Zhang S, Iwata K, Lachenmann MJ, Peng JW, Li S, Stimson ER, et al. The Alzheimer's peptide Aβ adopts a collapsed coil structure in water. J Struct Biol. 2000; 130: 130–141. https://doi.org/10.1006/jsbi.2000.4288 PMID: 10940221

186. Vitalis A, Caflisch A. Micelle-like architecture of the monomer ensemble of Alzheimer's amyloid-β peptide in aqueous solution and its implications for Aβ aggregation. J Mol Biol. 2010; 403: 148–165. https://doi.org/10.1016/j.jmb.2010.08.003 PMID: 20709081

187. Nag S, Sarkar B, Bandyopadhyay A, Sahoo B, Sreenivasan VAK, Kombrabail M, et al. Nature of amyloid-β monomer and the monomer-oligomer equilibrium. J Biol Chem. 2011; 286: 13827–13833. https://doi.org/10.1074/jbc.M110.199885 PMID: 21349839

188. Zhang-Hagen B, Biehl R, Nagel-Steger L, Radulescu A, Richter D, Willbold D. Monomeric amyloid beta peptide in hexafluoroisopropanol detected by small angle neutron scattering. PLoS ONE. 2016; 11: e0150267. https://doi.org/10.1371/journal.pone.0150267 PMID: 26919121

189. Roche J, Shen Y, Lee JH, Ying J, Bax A. Monomeric Aβ$^{1–40}$ and Aβ$^{1–42}$ peptides in solution adopt very similar Ramachandran map distribution that closely resemble random coil. Biochemistry. 2016; 55: 762–775. https://doi.org/10.1021/acs.biochem.5b01259 PMID: 26780756

190. Riek R, Güntert P, Döbeli H, Wipf B, Wüthrich K. NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, Aβ(1–40)^{ox} and Aβ(1–42)^{ox}. Eur J Biochem. 2001; 268: 5930–5936. PMID: 11722581

191. Vivekanandan S, Brender JR, Lee SY, Ramamoorthy A. A partially folded structure of amyloid-beta(1–40) in an aqueous environment. Biochem Biophys Res Comm. 2011: 411: 312–316. https://doi.org/10.1016/j.bbrc.2011.06.133 PMID: 21726530

192. Ball KA, Phillips AH, Wemmer DE, Head-Gordon T. Differences in β-strand populations of monomeric Aβ40 and Aβ42. Biophys J. 2013; 104: 2714–2724. https://doi.org/10.1016/j.bpj.2013.04.056 PMID: 23790380

193. Bhowmik D, MacLaughlin CM, Chandrakesan M, Ramesh P, Venkatramani R, Walker GC, et al. pH changes the aggregation propensity of amyloid-β without altering the monomer conformation. Chem Phys Phys Chem. 2014; 16: 885–889.

194. Hoyer W, Grönwall C, Jonsson A, Ståhl S, Härd T. Stabilization of a β-hairpin in monomeric Alzheimer's amyloid-β peptide inhibits amyloid formation. Proc Natl Acad Sci USA. 2008; 105: 5099–5104. https://doi.org/10.1073/pnas.0711731105 PMID: 18375754

195. Garzon-Rodriguez W, Sepulveda-Becerra M, Milton S, Glabe CG. Soluble amyloid Aβ-(1-40) exists as a stable dimer at low concentrations. J Biol Chem. 1997; 272: 21037–21044. PMID: 9261105

196. McDonald JM, Savva GM, Brayne C, Welzel AT, Forster G, Shankar GM, et al. The presence of sodium dodecyl sulphate-stable Aβ dimers is strongly associated with Alzheimer-type dementia. Brain. 2010; 133: 1328–1341. https://doi.org/10.1093/brain/awq065 PMID: 20403962

197. Villemagne VL, Perez KA, Pike KE, Kok WM, Rowe CC, White AR, et al. Blood-borne amyloid-β dimer correlates with clinical markers of Alzheimer's disease. J Neurosci. 2010; 30: 6315–6322. https://doi.org/10.1523/JNEUROSCI.5180-09.2010 PMID: 20445057

198. McDonald JM, O'Malley TT, Liu W, Mably AJ, Brinkmalm G, Portelius E, et al. The aqueous phase of Alzheimer's disease brain contains assemblies built from ~4 and ~7 kDa Aβ species. Alzheimer's Dement. 2015; 11: 1286–1305.

199. Müller-Schiffmann A, Herring A, Abdel-Hafiz L, Chepkova AN, Schäble S, Wendel D, et al. Amyloid-β dimers in the absence of plaque pathology impair learning and synaptic plasticity. Brain. 2016; 139: 509–525. https://doi.org/10.1093/brain/awv355 PMID: 26657517

200. Ryan TM, Roberts BR, McColl G, Hare DJ, Doble PA, Lin QX, et al. Stabilization of nontoxic Aβ-oligomers: insights into the mechanism of action of hydroxyquinolines in Alzheimer's disease. J Neurosci. 2015; 35: 2871–2884. https://doi.org/10.1523/JNEUROSCI.2912-14.2015 PMID: 25698727

201. Huang D, Zimmerman MI, Martin PK, Nix AJ, Rosenberry TL, Paravastu AK. Antiparallel β-sheet structure within the C-terminal region of 42-residue Alzheimer's amyloid-β peptides when they form 150 kDa oligomers. J Mol Biol. 2015; 427: 2319–2328. https://doi.org/10.1016/j.jmb.2015.04.004 PMID: 25889972

202. Tay WM, Huang D, Rosenberry TL, Paravastu AK. The Alzheimer's amyloid-β(1–42) peptide forms off-pathway oligomers and fibrils that are distinguished structurally by intermolecular organization. J Mol Biol. 2013; 425: 2494–2508. https://doi.org/10.1016/j.jmb.2013.04.003 PMID: 23583777

203. Garai K, Frieden C. Quantitative analysis of the time course of Aβ oligomerization and subsequent growth steps using tetramethylrhodamine-labeled Aβ. Proc Natl Acad Sci USA. 2013; 110: 3321–3326. https://doi.org/10.1073/pnas.1222478110 PMID: 23401512

204. Tiiman A, Krishtal J, Palumaa P, Tõugu V. In vitro fibrillization of Alzheimer's amyloid-β peptide (1–42). AIP Advances. 2015; 5: 092401.

205. Ono K, Condron MC, Teplow DB. Effects of the English (H6R) and Tottori (D7N) familial Alzheimer disease mutations on amyloid β-protein assembly and toxicity. J Biol Chem. 2010; 285: 23186–23197. https://doi.org/10.1074/jbc.M109.086496 PMID: 20452980

206. Lin MY, Liu JS, Wu YS, Peng CH, Chang YY. A novel *APP* mutation (D678H) in a Taiwanese patient exhibiting dementia and cerebral microvasculopathy. J Clin Neurosci. 2014; 21: 513–515. PMID: 23931937

207. Benilova I, Gallardo R, Ungureanu AA, Castillo Cano V, Snellinx A, Ramakers M, et al. The Alzheimer disease protective mutation A2T modulates kinetic and thermodynamic properties of amyloid-β (Aβ) aggregation. J Biol Chem. 2014; 289: 30977–30989. https://doi.org/10.1074/jbc.M114.599027 PMID: 25253695

208. Messa M, Colombo L, del Favero E, Cantu L, Stoilova T, Cagnotto A, et al. The peculiar role of the A2V mutation in amyloid-β (Aβ) 1–42 molecular assembly. J Biol Chem. 2014; 289: 24143–24152. https://doi.org/10.1074/jbc.M114.576256 PMID: 25037228

209. Zheng X, Liu D, Roychaudhuri R, Teplow DB, Bowers MT. Amyloid β-protein assembly: differential effects of the protective A2T mutation and recessive A2V familial Alzheimer's disease mutation. ACS Chem. Neurosci. 2015; 6: 1732–1740.

210. Bitan G, Vollers SS, Teplow DB. Elucidation of primary structure elements controlling early amyloid β-protein oligomerization. J Biol Chem. 2003; 278: 34882–34889. https://doi.org/10.1074/jbc.M300825200 PMID: 12840029

211. Christopeit T, Hortschansky P, Schroeckh V, Gührs K, Zandomeneghi G, Fändrich M. Mutagenic analysis of the nucleation propensity of oxidized Alzheimer's β-amyloid peptide. Protein Sci. 2005; 14: 2125–2131. https://doi.org/10.1110/ps.051470405 PMID: 15987892

212. Armstrong AH, Chen J, Fortner McKoy A, Hecht MH. Mutations that replace aromatic side chains promote aggregation of the Alzheimer's Aβ peptide. Biochemistry. 2011; 50: 4058–4067. https://doi.org/10.1021/bi200268w PMID: 21513285

213. Chen YR, Huang HB, Lo CJ, Wang CC, Ho LK, Liu HT, et al. Effect of alanine replacement of L17 and F19 on the aggregation and neurotoxicity of Arctic type Aβ_{40}. PLoS ONE. 2013, 8: e61874. https://doi.org/10.1371/journal.pone.0061874 PMID: 23634215

214. Adler J, Scheidt HA, Krüger M, Thomas L, Huster D. Local interactions influence the fibrillation kinetics, structure and dynamics of Aβ(1–40) but leave the general fibril structure unchanged. Phys Chem Chem Phys. 2014; 16: 7461–7471. https://doi.org/10.1039/c3cp54501f PMID: 24626742

215. Yagi-Utsumi M, Dobson CM. Conformational effects of the A21G Flemish mutation on the aggregation of amyloid β peptide. Bio Pharm Bull. 2015; 38: 1668–1672.

216. Kim W, Hecht MH. Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide. Proc Natl Acad Sci USA. 2006; 103: 15824–15829. https://doi.org/10.1073/pnas.0605629103 PMID: 17038501

217. Lv Z, Roychaudhuri R, Condron MC, Teplow DB, Lyubchenko Y. Mechanism of amyloid β-protein dimerization determined using single-molecule AFM force spectroscopy. Sci Rep. 2013; 3: 2880. https://doi.org/10.1038/srep02880 PMID: 24096987

218. Chemuru S, Kodali R, Wetzel R. C-Terminal threonine reduces Aβ_{43} amyloidogenicity compared with Aβ_{42}. J Mol Biol. 2016; 428: 274–291. https://doi.org/10.1016/j.jmb.2015.06.008 PMID: 26122432

219. Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB. Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways. Proc Natl Acad Sci USA. 2003; 100: 330–335. https://doi.org/10.1073/pnas.222681699 PMID: 12506200

220. Lambermon MH, Rappaport RV, McLaurin J. Biophysical characterization of longer forms of amyloid beta peptides: possible contribution to flocculent plaque formation. J Neurochem. 2005; 95: 1667–1676. https://doi.org/10.1111/j.1471-4159.2005.03497.x PMID: 16300644

221. Saito T, Suemoto T, Brouwers N, Sleegers K, Funamoto S, Mihira N, et al. Potent amyloidogenicity and pathogenicity of Aβ43. Nat Neurosci. 2011; 14: 1023–1032. https://doi.org/10.1038/nn.2858 PMID: 21725313

222. Vandersteen A, Hubin E, Sarroukh R, De Baets G, Schymkowitz J, Rousseau F, et al. A comparative analysis of the aggregation behaviour of amyloid-β peptide variants. FEBS Lett. 2012; 586: 4088–4093. https://doi.org/10.1016/j.febslet.2012.10.022 PMID: 23103738

223. Zou K, Liu J, Watanabe A, Hiraga S, Liu S, Tanabe C, et al. Aβ_{43} is the earliest-depositing Aβ species in APP transgenic mouse brain and is converted to Aβ_{41} by two active domains of ACE. Am J Pathol. 2013; 182: 2322–2331. https://doi.org/10.1016/j.ajpath.2013.01.053 PMID: 23562443

224. Sandebring A, Welander H, Winblat B, Graff C, Tjernberg LO. The pathogenic Aβ43 is enriched in familial and sporadic Alzheimer disease. PLoS ONE. 2013; 8: e55847. https://doi.org/10.1371/journal.pone.0055847 PMID: 23409063

225. Conicella AE, Fawzi NL. The C-terminal threonine of Aβ43 nucleates toxic aggregation via structural and dynamic changes in monomers and protofibrils. Biochemistry. 2014; 53: 3095–3107. https://doi.org/10.1021/bi500131a PMID: 24773532

226. Chemuru S, Kodali R, Wetzel R. Improved chemical synthesis of hydrophobic Aβ peptides using addition of C-terminal lysines later removed by carboxypeptidase B. Biopolymers. 2014; 102: 206–221. https://doi.org/10.1002/bip.22470 PMID: 24488729

227. Burnouf S, Gorsky MK, Dols J, Grönke S, Partridge L. Aβ_{43} is neurotoxic and primes aggregation of Aβ_{40} in vivo. Acta Neuropathol. 2015; 130: 35–47. https://doi.org/10.1007/s00401-015-1419-y PMID: 25862636

228. Kretner B, Trambauer J, Fukumori A, Mielke J, Kuhn PH, Kremmer E, et al. Generation and deposition of Aβ43 by the virtually inactive presenilin-1 L435F mutant contradicts the presenilin loss-of-function hypothesis of Alzheimer's disease. EMBO Mol Med. 2016; 8: 458–465. https://doi.org/10.15252/emmm.201505952 PMID: 26988102

229. Soreghan B, Kosmoski J, Glabe C. Surfactant properties of Alzheimer's Aβ peptides and the mechanism of amyloid aggregation. J Biol Chem. 1994; 269: 28551–28554. PMID: 7961799

230. Yong W, Lomakin A, Kirkitadze MD, Teplow DB, Chen SH, Benedek GB. Structure determination of micelle-like intermediates in amyloid β-protein fibril assembly by using small angle neutron scattering. Proc Natl Acad Sci USA. 2002; 99: 150–154. https://doi.org/10.1073/pnas.012584899 PMID: 11756677

231. Aulikki Wälti M, Orts J, Vögeli B, Campioni S, Riek R. Solution NMR studies of recombinant Aβ(1–42): from the presence of a micellar entity to residual β-sheet structure in the soluble species. ChemBioChem. 2015; 16: 659–669. https://doi.org/10.1002/cbic.201402595 PMID: 25676345

232. Kirkitadze MD, Condron MM, Teplow DB. Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis. J Mol Biol. 2001; 312: 1103–1119. https://doi.org/10.1006/jmbi.2001.4970 PMID: 11580253

233. Fezoui Y, Teplow DB. Kinetic studies of amyloid β-protein fibril assembly. J Biol Chem. 2002; 277: 36948–36954. https://doi.org/10.1074/jbc.M204168200 PMID: 12149256

234. Sciaretta KL, Gordon DJ, Petkova AT, Tycko R. Aβ 40-Lactam(D23/K28) models a conformation highly favourable for nucleation of amyloid. Biochemistry. 2005; 44: 6003–6014. https://doi.org/10.1021/bi0474867 PMID: 15835889

235. Dubnovitsky A, Sandberg A, Rahman MM, Benilova I, Lendel C, Härd T. Amyloid-β protofibrils: size, morphology and synaptotoxicity of an engineered mimic. PLoS ONE. 2013; 8: e66101. https://doi.org/10.1371/journal.pone.0066101 PMID: 23843949

236. Tew DJ, Bottomley SP, Smith DP, Ciccotosto GD, Babon J, Hinds MG, et al. Stabilization of neurotoxic soluble β-sheet-rich conformations of the Alzheimer's disease amyloid-β peptide. Biophys J. 2008; 94: 2752–2766. https://doi.org/10.1529/biophysj.107.119909 PMID: 18065467

237. Necula M, Kayed R, Milton S, Glabe CG. Small molecule inhibitors of aggregation indicate that amyloid β oligomerization and fibrillization pathways are independent and distinct. J Biol Chem. 2007; 282: 10311–10324. https://doi.org/10.1074/jbc.M608207200 PMID: 17284452

238. Ryan TM, Friedhuber A, Lind M, Howlett GJ, Masters C, Roberts BR. Small amphipathic molecules modulate secondary structure and amyloid fibril-forming kinetics of Alzheimer disease peptide $Aβ_{1-42}$. J Biol Chem. 2012; 287: 16947–16954. https://doi.org/10.1074/jbc.M111.321778 PMID: 22461629

239. Hashemi Shabestari M, Meeuwenoord NJ, Filipov DV, Huber M. Interaction of the amyloid β peptide with sodium dodecyl sulphate as a membrane-mimicking detergent. J Biol Phys. 2016; 42: 299–315. https://doi.org/10.1007/s10867-016-9408-5 PMID: 26984615

240. Habchi J, Arosio P, Perni M, Costa AR, Yagi-Utsumi M, Joshi P, et al. An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic Aβ42 aggregates linked with Alzheimer's disease. Sci Adv. 2016; 2: e1501244. https://doi.org/10.1126/sciadv.1501244 PMID: 26933687

241. Luo J, Wärmländer SKTS, Gräslund A, Abrahams JP. Non-chaperone proteins can inhibit aggregation and cytotoxicity of Alzheimer amyloid β peptide. J Biol Chem. 2014; 289: 27766–27775. https://doi.org/10.1074/jbc.M114.574947 PMID: 25100721

242. Ono K, Li L, Takamura Y, Yoshiike Y, Zhu L, Han F, et al. Phenolic compounds prevent amyloid β-protein oligomerization and synaptic dysfunction by site-specific binding. J Biol Chem. 2012; 287: 14631–14643. https://doi.org/10.1074/jbc.M111.325456 PMID: 22393064

243. Zheng X, Gessel MM, Wisniewski ML, Visvanathan K, Wright DL, Bahr BA, et al. Z-Phe-Ala-diazomethylketone (PADK) disrupts and remodels early oligomer states of the Alzheimer disease Aβ42 protein. J Biol Chem. 2012; 287: 6084–6088. https://doi.org/10.1074/jbc.C111.328575 PMID: 22253440

244. Lopez del Amo JM, Fink U, Dasari M, Grelle G, Wanker EE, Bieschke J, et al. Structural properties of EGCG-induced, nontoxic Alzheimer's disease Aβ oligomers. J Mol Biol. 2012; 421: 517–524. https://doi.org/10.1016/j.jmb.2012.01.013 PMID: 22300765

245. Nerelius C, Sandegren A, Sargasyan H, Raunak R, Leijonmarck H, Chatterjee U, et al. α-Helix targeting reduced amyloid-β peptide toxicity. Proc Natl Acad Sci USA. 2009; 106: 9191–9196. https://doi.org/10.1073/pnas.0810364106 PMID: 19458258

246. Hamley IW. The amyloid beta peptide: a chemist's perspective. Role in Alzheimer's and fibrillization. Chem Rev. 2012; 112: 5147–5192. https://doi.org/10.1021/cr3000994 PMID: 22813427

247. Choo-Smith LP, Surewicz WK. The interaction between Alzheimer amyloid β(1–40) peptide and ganglioside G_{M1}-containing membranes. FEBS Lett. 1997; 402: 95–98. PMID: 9037173

248. Kakio A, Nishimoto S, Yanagisawa K, Kozutsumi Y, Matsuzaki K. Cholesterol-dependent formation of GM1 ganglioside-bound amyloid β-protein, an endogenous seed for Alzheimer amyloid. J Biol Chem. 2001; 276: 24985–24990. https://doi.org/10.1074/jbc.M100252200 PMID: 11342534

249. Williamson MP, Suzuki Y, Bourne NT, Asakura T. Binding of amyloid β-peptide to ganglioside micelles is dependent on histidine-13. Biochem J. 2006; 397: 483–490. https://doi.org/10.1042/BJ20060293 PMID: 16626304

250. Zhang YJ, Shi JM, Bai CJ, Wang H, Li HY, Wu Y, et al. Intra-membrane oligomerization and extra-membrane oligomerization of amyloid-β peptide are competing processes as a result of distinct patterns of motif interplay. J Biol Chem. 2012; 287: 748–756. https://doi.org/10.1074/jbc.M111.281295 PMID: 22105077

251. Fukunaga S, Ueno H, Yamaguchi T, Yano Y, Hoshino M, Matsuzaki K. GM1 cluster mediates formation of toxic Aβ fibrils by providing hydrophobic environments. Biochemistry. 2012; 51: 8125–8131. https://doi.org/10.1021/bi300839u PMID: 23009396

252. Ueno H, Yamaguchi T, Fukunaga S, Okada Y, Yano Y, Hoshino M, Matsuzaki K. Comparison between the aggregation of human and rodent amyloid β-proteins in GM1 ganglioside clusters. Biochemistry. 2014; 53: 7523–7530. https://doi.org/10.1021/bi501239q PMID: 25399921

253. Matsuzaki K. How do membranes initiate Alzheimer's disease? Formation of toxic amyloid fibrils by the amyloid β-protein on gangliosides clusters. Acc Chem Res. 2014; 47: 2397–2404. https://doi.org/10.1021/ar500127z PMID: 25029558

254. Yi X, Zhang Y, Gong M, Yu X, Darabedian N, Zheng J, et al. Ca^{2+} interacts with Glu-22 of Aβ(1–42) and phospholipid bilayers to accelerate the Aβ(1–42) aggregation below the critical micelle concentration. Biochemistry. 2015; 54: 6323–6332. https://doi.org/10.1021/acs.biochem.5b00719 PMID: 26426973

255. Lambert MP, Barlow AK, Chromy BA, Edwards C, Freed R, Liosatos M, et al. Diffusible, nonfibrillar ligands derived from Aβ1–42 are potent central nervous system neurotoxins. Proc Natl Acad Sci USA. 1998; 95: 6448–6453. PMID: 9600986

256. Dahlgren KN, Manelli AM, Stine WB Jr, Baker LK, Krafft GA, LaDu MJ. Oligomeric and fibrillar species of amyloid-β peptides differentially affect neuronal viability. J Biol Chem. 2002; 277: 32046–32053. https://doi.org/10.1074/jbc.M201750200 PMID: 12058030

257. Glabe CG. Structural classification of toxic amyloid oligomers. J Biol Chem. 2008; 283: 29639–29643. https://doi.org/10.1074/jbc.R800016200 PMID: 18723507

258. Gellermann GP, Byrnes H, Striebinger A, Ullrich K, Mueller R, Hillen H, et al. Aβ-globulomers are formed independently of the fibril pathway. Neurobiol Dis. 2008; 30: 212–220. https://doi.org/10.1016/j.nbd.2008.01.010 PMID: 18353662

259. Matsumura S, Shinoda K, Yamada M, Yokojima S, Inoue M, Ohnishi T, et al. Two distinct amyloid β-protein (Aβ) assembly pathways leading to oligomers and fibrils identified by combined fluorescence correlation spectroscopy, morphology and toxicity analyses. J Biol Chem. 2011; 286: 11555–11562. https://doi.org/10.1074/jbc.M110.181313 PMID: 21292768

260. Kumar A, Bullard RL, Paslay LC, Singh D, Bienkiewicz EA, Morgan SE, Rangachari V. Non-esterified fatty acids generate distinct low-molecular weight amyloid-β (Aβ42) oligomers along pathway different from fibril formation. PLoS ONE. 2011; 6: e18759. https://doi.org/10.1371/journal.pone.0018759 PMID: 21526230

261. Vignaud H, Bobo C, Lascu I, Sörgjerd KM, Zako T, Maeda M, et al. A structure-toxicity study of Aβ$_{42}$ reveals a new antiparallel aggregation pathway. PLoS ONE. 2013; 8: e80262. https://doi.org/10.1371/journal.pone.0080262 PMID: 24244667

262. Liu P, Reed MN, Kotilinek LA, Grant MKO, Forster CL, Qiang W, et al. Quaternary structure defines a large class of amyloid-β oligomers neutralized by sequestration. Cell Reports. 2015; 11: 1760–1771. https://doi.org/10.1016/j.celrep.2015.05.021 PMID: 26051935

263. Stine WB Jr, Dahlgren KN, Krafft GA, LaDu MJ. In vitro characterization of conditions for amyloid-β peptide oligomerization and fibrillization. J Biol Chem. 2003; 278: 11612–11622. https://doi.org/10.1074/jbc.M210207200 PMID: 12499373

264. Bernstein SL, Dupuis NF, Lazo ND, Wyttenbach T, Condron MM, Bitan G, et al. Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aethiology of Alzheimer's disease. Nat Chem. 2009; 1: 326–331. https://doi.org/10.1038/nchem.247 PMID: 20703363

265. Fu F, Aucoin D, Davis J, Van Nostrand WE, Smith SO. Mechanism of nucleated conformational conversion of Aβ42. Biochemistry. 2015; 54: 4197–4207. https://doi.org/10.1021/acs.biochem.5b00467 PMID: 26069943

266. Ahmed M, Davies J, Aucoin D, Sato T, Ahuja S, Aimoto S, et al. Structural conversions of neurotoxic amyloid-β$_{1–42}$ oligomers to fibrils. Nat Struct Mol Biol. 2010; 17: 561–568. https://doi.org/10.1038/nsmb.1799 PMID: 20383142

267. Mastrangelo IA, Ahmed M, Sato T, Liu W, Wang C, Hough P, et al. High-resolution atomic force microscopy of soluble Aβ42 oligomers. J Mol Biol. 2006; 358: 106–119. https://doi.org/10.1016/j.jmb.2006.01.042 PMID: 16499926

268. Wolff M, Unuchek D, Zhang B, Gordeliy V, Willibold D, Nagel-Steger L. Amyloid β oligomeric species present in the lag phase of amyloid formation. PLoS ONE. 2015; 10: e0127865. https://doi.org/10.1371/journal.pone.0127865 PMID: 26024352

269. Kusumoto Y, Lomakin A, Teplow DB, Benedek GB. Temperature dependence of amyloid β-protein fibrillization. Proc Natl Acad Sci USA. 1998; 95: 12277–12282. PMID: 9770477

270. Lee J, Culyba EK, Powers ET, Kelly JW. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. Nat Chem Biol. 2011; 7: 602–609. https://doi.org/10.1038/nchembio.624 PMID: 21804535

271. Lashuel HA, Hartley D, Petre BM, Waltz T, Lansbury PT Jr. Neurodegenerative disease: amyloid pores from pathogenic mutations. Nature. 2002; 418: 291.

272. Luo J, Wärmländer SKTS, Gräslund A, Abrahams JP. Alzheimer peptides aggregate into transient nanoglobules that nucleate fibrils. Biochemistry. 2014; 53: 6302–6308. https://doi.org/10.1021/bi5003579 PMID: 25198136

273. Economou NJ, Giammona MJ, Do TD, Zheng X, Teplow DB, Buratto SK, et al. Amyloid β-protein assembly and Alzheimer's disease: dodecamers of Aβ42, but not of Aβ40, seed fibril formation. J Am Chem Soc. 2016; 138: 1772–1775. https://doi.org/10.1021/jacs.5b11913 PMID: 26839237

274. Chimon S, Shaibat MS, Jones CR, Calero DC, Aizezi B, Ishii Y. Evidence of fibril-like β-sheet structure in a neurotoxic amyloid intermediate of Alzheimer's β-amyloid. Nat Struct Mol Biol. 2007; 14: 1157–1164. https://doi.org/10.1038/nsmb1345 PMID: 18059284

275. Parthasarathy S, Inoue M, Xiao Y, Matsumura Y, Nabeshima Y, Hoshi M, et al. Structural insight into an Alzheimer's brain-derived spherical assembly of amyloid β by solid-state NMR. J Am Chem Soc. 2015; 137: 6480–6483. https://doi.org/10.1021/jacs.5b03373 PMID: 25938164

276. Ladiwala ARA, Dordick JS, Tessier PM. Aromatic small molecules remodel toxic soluble oligomers of amyloid β through three independent pathways. J Biol Chem. 2011; 286: 3209–3218. https://doi.org/10.1074/jbc.M110.173856 PMID: 21098486

277. Williams AD, Sega MD, Chen M, Kheterpal I, Geva M, Berthelier V, et al. Structural properties of Aβ protofibrils stabilized by a small molecule. Proc Natl Acad Sci. USA. 2005; 102: 7115–7120. https://doi.org/10.1073/pnas.0408582102 PMID: 15883377

278. Bieschke J, Herbst M, Wiglenda T, Friedrich RP, Boeddrich A, Schiele F, et al. Small-molecule conversion of toxic oligomers to non-toxic β-sheet-rich amyloid fibrils. Nat Chem Biol. 2012; 8: 93–101.

279. Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB. Amyloid β-protein fibrillogenesis. J Biol Chem. 1997; 272: 22364–22372.

280. Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, Lomakin A, et al. Amyloid β-protein fibrillogenesis. J Biol Chem. 1999; 274: 25945–25952.

281. Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, Forsell C, et al. The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibrils formation. Nat Neurosci. 2001; 4: 887–893. https://doi.org/10.1038/nn0901-887 PMID: 11528419

282. Doi T, Masuda Y, Irie K, Akagi KI, Imazawa T, Takegoshi K. Solid-state NMR analysis of the β-strand orientation of the protofibrils of amyloid β-protein. Biochem Biophys Res Comm. 2012; 428: 458–462. https://doi.org/10.1016/j.bbrc.2012.10.096 PMID: 23131555

283. Schmidt M, Rohou A, Lasker K, Yadav JK, Schiene-Fischer C, Fändrich M, et al. Peptide dimer structure in an Aβ(1–42) fibril visualized with cryo-EM. Proc Natl Acad Sci USA. 2015; 112: 11858–11863. https://doi.org/10.1073/pnas.1503455112 PMID: 26351699

284. Sachse C, Fändrich M, Grigorieff N. Paired β-structure of an Aβ(1–40) amyloid fibril revealed by electron microscopy. Proc Natl Acad Sci USA. 2008; 105: 7462–7466. https://doi.org/10.1073/pnas.0712290105 PMID: 18483195

285. Cohen SIA, Linse S, Luheshi LM, Hellstrand E, White DA, Rajah L, et al. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. Proc Natl Acad Sci USA. 2013; 110: 9758–9763. https://doi.org/10.1073/pnas.1218402110 PMID: 23703910

286. Jeong JS, Ansaloni A, Mezzenga R, Lashuel HA, Dietler G. Novel mechanistic insight into the molecular basis of amyloid polymorphism and secondary nucleation during amyloid formation. J Mol Biol. 2013; 425: 1765–1781. https://doi.org/10.1016/j.jmb.2013.02.005 PMID: 23415897

287. Xiao Y, Ma B, McElheny D, Parthasarathy S, Long F, Hoshi M, et al. Aβ(1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. Nat Struct Mol Biol. 2015; 22: 499–505. https://doi.org/10.1038/nsmb.2991 PMID: 25938662

288. Tycko R. Amyloid polymorphism: structural basis and neurobiological relevance. Neuron. 2015; 86: 632–645. https://doi.org/10.1016/j.neuron.2015.03.017 PMID: 25950632

289. Cerf E, Sarroukh R, Tamamizu-Kato S, Breydo L, Derclaye S, Dufrêne YF, et al. Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. Biochem J. 2009; 421: 415–423. https://doi.org/10.1042/BJ20090379 PMID: 19435461

290. Gu L, Liu C, Stroud JC, Ngo S, Jiang L, Guo Z. Antiparallel triple-strand architecture for prefibrillar Aβ42 oligomers. J Biol Chem. 2014; 289: 27300–27313. https://doi.org/10.1074/jbc.M114.569004 PMID: 25118290

291. Gu L, Liu C, Guo Z. Structural insight into Aβ42 oligomers using site-directed spin labelling. J Biol Chem. 2013; 288: 18673–18683. https://doi.org/10.1074/jbc.M113.457739 PMID: 23687299

292. Qiang W, Yau WM, Luo Y, Mattson MP, Tycko R. Antiparallel β-sheet architecture in Iowa-mutant β-amyloid fibrils. Proc Natl Acad Sci USA. 2012; 109: 4443–4448. https://doi.org/10.1073/pnas.1111305109 PMID: 22403062

293. Hubin E, Deroo S, Kaminski Schierle G, Kaminski C, Serpell L, Subramaniam V, et al. Two distinct β-sheet structures in Italian-mutant amyloid-beta fibrils: a potential link to different clinical phenotypes. Cell Mol Life Sci. 2015; 72: 4899–4913. https://doi.org/10.1007/s00018-015-1983-2 PMID: 26190022

294. Barghorn S, Nimmrich V, Striebinger A, Krantz C, Keller P, Janson B, et al. Globular amyloid β-peptide$_{1-42}$ oligomer—a homogenous and stable neuropathological protein in Alzheimer's disease. J Neurochem. 2005; 95: 834–847. https://doi.org/10.1111/j.1471-4159.2005.03407.x PMID: 16135089

295. Lesné SE, Koh MT, Kotilinek L, Kayed R, Glabe CG, et al. A specific amyloid-beta protein assembly in the brain impairs memory. Nature. 2006; 440: 352–357. https://doi.org/10.1038/nature04533 PMID: 16541076

296. Lesné SE, Sherman MA, Grant M, Kuskowski M, Schneider JA, Bennett DA, et al. Brain amyloid-β oligomers in ageing and Alzheimer's disease. Brain. 2013; 136: 1383–1398. https://doi.org/10.1093/brain/awt062 PMID: 23576130

297. Kumar A, Paslay LC, Lyons D, Morgan SE, Correia JJ, Rangachari V. Specific soluble oligomers of amyloid-β peptide undergo replication and form non-fibrillar aggregates in interfacial environments. J Biol Chem. 2012; 287: 21253–21264. https://doi.org/10.1074/jbc.M112.355156 PMID: 22544746

298. Kumar A, Pate KM, Moss MA, Dean DE, Rangachari V. Self-propagative replication of Aβ oligomers suggests potential transmissibility in Alzheimer disease. PLoS ONE. 2014; 9: e111492. https://doi.org/10.1371/journal.pone.0111492 PMID: 25365422

299. Dean DE, Pate KM, Moss MA, Rangachari V. Conformational dynamics of specific Aβ oligomers govern their ability to replicate and induce neuronal apoptosis. Biochemistry. 2016; 55: 2238–2250. https://doi.org/10.1021/acs.biochem.6b00161 PMID: 27013020

300. Sebollela A, Mustata GM, Luo K, Velasco PT, Viola KL, Cline EN, et al. Elucidating molecular mass and shape of a neurotoxic Aβ oligomer. ACS Chem Neurosci. 2014; 5: 1238–1245. https://doi.org/10.1021/cn500156r PMID: 25343357

301. Ladiwala ARA, Litt J, Kane RS, Aucoin DS, Smith SO, Ranjan S, et al. Conformational differences between two amyloid β oligomers of similar size and dissimilar toxicity. J Biol Chem. 2012; 287: 24765–24773. https://doi.org/10.1074/jbc.M111.329763 PMID: 22547072

302. Zhang H, Li Y, Xu X, Sun T, Fuchs H, Chi L. Ion strength and pH sensitive phase transition of N-isobutyryl-L-(D)-cysteine monolayers on Au(111) surfaces. Langmuir. 2010; 26: 7343–7348. https://doi.org/10.1021/la904237d PMID: 20355723

303. Ramachandran G, Udgaonkar JB. Mechanistic studies unravel the complexity inherent in tau aggregation leading to Alzheimer's disease and the tauopathies. Biochemistry. 2013; 52: 4107–4126. https://doi.org/10.1021/bi400209z PMID: 23721410

304. Wang Y, Mandelkow E. Tau in physiology and pathology. Nat Rev Neurosci. 2016; 17: 5–21. https://doi.org/10.1038/nrn.2015.1 PMID: 26631930

305. Barré P, Eliezer D. Folding of the repeat domain of tau upon binding to lipid surfaces. J Mol Biol. 2006; 362: 312–326. https://doi.org/10.1016/j.jmb.2006.07.018 PMID: 16908029

306. Georgieva ER, Xiao S, Borbat PP, Freed JH, Eliezer D. Tau binds to lipid membrane surfaces via short amphipathic helices located in its microtubule-binding repeats. Biophys J. 2014; 107: 1441–1452. https://doi.org/10.1016/j.bpj.2014.07.046 PMID: 25229151

307. Kadavath H, Jaremko M, Jaremko L, Biernat J, Mandelkow E, Zweckstetter M. Folding of the tau protein on microtubules. Angew Chem Int Ed. 2015; 54: 10347–10351.

308. Li XH, Culver JA, Rhoades E. Tau binds to multiple tubulin dimers with helical structures. J Am Chem Soc. 2015; 137: 9218–9221. https://doi.org/10.1021/jacs.5b04561 PMID: 26165802

309. Grüning CSR, Mirecka EA, Klein AN, Mandelkow E, Willbold D, Marino SF, et al. Alternative conformations of the tau repeat domain in complex with an engineered binding protein. J Biol Chem. 2014; 289: 23209–23218. https://doi.org/10.1074/jbc.M114.560920 PMID: 24966331

310. Andronesi OC, von Bergen M, Biernat J, Seidel K, Griesinger C, Mandelkow E, et al. Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. J Am Chem Soc. 2008; 130: 5922–5928. https://doi.org/10.1021/ja7100517 PMID: 18386894

311. Friedhoff P, von Bergen M, Mandelkow EM, Davies P, Mandelkow E. A nucleated assembly mechanism of Alzheimer paired helical filaments. Proc Natl Acad Sci USA. 1998; 95: 15712–15717. PMID: 9861035

312. Congdon EE, Kim S, Bonchak J, Songrug T, Matzavinos A, Kurett J. Nucleation-dependent tau filament formation. J Biol Chem. 2008; 283: 13806–13816. https://doi.org/10.1074/jbc.M800247200 PMID: 18359772

313. Ramachandran G, Udgaonkar JB. Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by tau protein. J Biol Chem. 2011; 286: 38948–38959. https://doi.org/10.1074/jbc.M111.271874 PMID: 21931162

314. Ramachandran G, Milán-Garcés EA, Udgaonkar JB, Puranik M. Resonance Raman spectroscopic measurements delineate the structural changes that occur during tau fibril formation. Biochemistry. 2014; 53: 6550–6565. https://doi.org/10.1021/bi500528x PMID: 25284680

315. Daebel V, Chinnathambi S, Biernat J, Schwalbe M, Habenstein B, Loquet A, et al. β-Sheet core of tau paired helical filaments revealed by solid-state NMR. J Am Chem Soc. 2012; 134: 13982–13989. https://doi.org/10.1021/ja305470p PMID: 22862303

316. Meyer V, Holden MR, Weismiller HA, Eaton GR, Eaton SS, Margittai M. Fracture and growth are competing forces determining the fate of conformers in tau fibril populations. J Biol Chem. 2016; 291: 12271–12281. https://doi.org/10.1074/jbc.M116.715557 PMID: 27080260

317. Wang C, Zhao C, Li D, Tian Z, Lai Y, Liu C. Versatile structures of α-synuclein. Front Mol Neurosci. 2016; 9: 48. https://doi.org/10.3389/fnmol.2016.00048 PMID: 27378848

318. Wang W, Perovic I, Chittuluru J, Kaganovich A, Nguyen LTT, Liao J, et al. A soluble α-synuclein construct forms a dynamic tetramer. Proc Natl Acad Sci USA. 2011; 108: 17797–17802. https://doi.org/10.1073/pnas.1113260108 PMID: 22006323

319. Bartels T, Choi JG, Selkoe DJ. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. Nature. 2011; 477: 107–110. https://doi.org/10.1038/nature10324 PMID: 21841800

320. Fauvet B, Mbefo MK, Fares MB, Desobry C, Michael S, Ardah MT, et al. α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and Escherichia coli exists predominantly as disordered monomer. J Biol Chem. 2012; 287: 15345–15364. https://doi.org/10.1074/jbc.M111.318949 PMID: 22315227

321. Moriarty GM, Janowska MK, Kang L, Baum J. Exploring the accessible conformations of N-terminal acetylated α-synuclein. FEBS Lett. 2013; 587: 1128–1138. https://doi.org/10.1016/j.febslet.2013.02.049 PMID: 23499431

322. Ulmer TS, Bax A, Cole NB, Nussbaum RL. Structure and dynamics of micelle-bound human α-synuclein. J Biol Chem. 2005; 280: 9595–9603. https://doi.org/10.1074/jbc.M411805200 PMID: 15615727

323. Rao JN, Jao CC, Hedge BG, Langen R, Ulmer TS. A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. J Am Chem Soc. 2010; 132: 8657–8668. https://doi.org/10.1021/ja100646t PMID: 20524659

324. Lashuel HA, Petre BM, Wall J, Simon M, Nowak RJ, Walz T, et al. α-Synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. J Mol Biol. 2002; 322: 1089–1102. PMID: 12367530

325. Chen SW, Drakulic S, Deas E, Ouberai M, Aprile FA, Arranz R, et al. Structural characterization of toxic oligomers that are kinetically trapped during α-synuclein fibril formation. Proc Natl Acad Sci USA. 2015; 112: E1994–E2003. https://doi.org/10.1073/pnas.1421204112 PMID: 25855634

326. Hayden EY, Kaur P, Williams TL, Matsui H, Yeh SR, Rousseau DL. Heme stabilization of α-synuclein oligomers during amyloid fibril formation. Biochemistry. 2015; 54: 4599–4610. https://doi.org/10.1021/acs.biochem.5b00280 PMID: 26161848

327. Celej MS, Sarroukh R, Goormaghtigh E, Fidelio GD, Ruysschaert JM, Raussens V. Toxic prefibrillar α-synuclein amyloid oligomers adopt a distinctive antiparallel β-sheet structure. Biochem J. 2012; 443: 719–726. https://doi.org/10.1042/BJ20111924 PMID: 22316405

328. Cremades N, Cohen SIA, Deas E, Abramov AY, Chen AY, Orte A, et al. Direct observation of the interconversion of normal and toxic forms of α-synuclein. Cell. 2012; 149: 1048–1059. https://doi.org/10.1016/j.cell.2012.03.037 PMID: 22632969

329. Gallea JI, Celej MS. Structural insights into amyloid oligomers of the Parkinson disease-related protein α-synuclein. J Biol Chem. 2014; 289: 26733–26742. https://doi.org/10.1074/jbc.M114.566695 PMID: 25143382

330. Paslawski W, Andreasen M, Bang Nielsen S, Lorenzen N, Thomsen K, Kaspersen JD, et al. High stability and cooperative unfolding of α-synuclein oligomers. Biochemistry. 2014; 53: 6252–6263. https://doi.org/10.1021/bi5007833 PMID: 25216651

331. Tossato L, Horrocks MH, Dear AJ, Knowles TPJ, Della Serra M, Cremades N, et al. Single-molecule FRET studies on alpha-synuclein oligomerization of Parkinson's disease genetically related mutants. Sci Rep. 2015; 5: 16696. https://doi.org/10.1038/srep16696 PMID: 26582456

332. Rodriguez JA, Ivanova MI, Sawaya MR, Cascio D, Reyes F, Shi D, et al. Structure of the toxic core of α-synuclein from invisible crystals. Nature. 2015; 525: 486–490. https://doi.org/10.1038/nature15368 PMID: 26352473

333. Tuttle MD, Comellas G, Nieuwkoop AJ, Covell DJ, Berthold DA, Kloepper KD, et al. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. Nat Struct Mol Biol. 2016; 23: 409–415. https://doi.org/10.1038/nsmb.3194 PMID: 27018801

334. Dearborn AD, Wall JS, Cheng N, Heymann JB, Kajava AV, Varkey J, et al. α-Synuclein amyloid fibrils with two entwined, asymmetrically associated, protofibrils. J Biol Chem. 2016; 291: 2310–2318. https://doi.org/10.1074/jbc.M115.698787 PMID: 26644467

335. Kundu B, Maiti NR, Jones EM, Surewicz KA, Vanik DL, Surewicz WK. Nucleation-dependent conformational conversion of the Y145Stop variant of human prion protein: structural clues for prion propagation. Proc Natl Acad Sci USA. 2003; 100: 12069–12074. https://doi.org/10.1073/pnas.2033281100 PMID: 14519851

336. Jones EM, Wu B, Surewicz K, Nadaud PS, Helmus JJ, Chen S, et al. Structural polymorphism in amyloids. New insights from studies with Y145Stop prion protein fibrils. J Biol Chem. 2011; 286: 42777–42784. https://doi.org/10.1074/jbc.M111.302539 PMID: 22002245

337. Helmus JJ, Surewicz K, Apostol MI, Surewicz WK, Jaroniec CP. Intermolecular alignment in Y145Stop human prion protein amyloid fibrils probed by solid-state NMR spectroscopy. J Am Chem Soc. 2011; 133: 2393–2403.

338. Honda RP, Yamaguchi K, Kuwata K. Acid-induced molten globule state of a prion protein. Crucial role of strand 1-helix1-strand 2 segment. J Biol Chem. 2014; 289: 30355–30363. https://doi.org/10.1074/jbc.M114.559450 PMID: 25217639

339. Gerber R, Tahiri-Alaoui A, Hove PJ, James W. Oligomerization of the human prion protein proceeds via a molten globule intermediate. J Biol Chem. 2007; 282: 6300–6307. https://doi.org/10.1074/jbc.M608926200 PMID: 17210575

340. Singh J, Udgaonkar JB. The pathogenic mutation T182A converts the prion protein into a molten globule-like conformation whose misfolding to oligomers but not to fibrils is drastically accelerated. Biochemistry. 2016; 55: 459–469. https://doi.org/10.1021/acs.biochem.5b01266 PMID: 26713717

341. Sabareesan AT, Udgaonkar JB. Pathogenic mutations within the disordered palindromic region of the prion protein induce structure therein and accelerate the formation of misfolded oligomers. J Mol Biol. 2016; 428: 3935–3947. https://doi.org/10.1016/j.jmb.2016.08.015 PMID: 27545411

342. Coleman BM, Harrison CF, Guo B, Masters CL, Barnham KJ, Lawson VA, et al. Pathogenic mutations within the hydrophobic domain of the prion protein lead to the formation of protease-sensitive prion species with increased lethality. J Virol. 2014; 88: 2690–2703. https://doi.org/10.1128/JVI.02720-13 PMID: 24352465

343. Sabareesan AT, Singh J, Roy S, Udgaonkar JB, Mathew MK, The pathogenic A116V mutation enhances ion-selective channel formation by prion protein in membranes. Biophys J. 2016; 110: 1766–1776. https://doi.org/10.1016/j.bpj.2016.03.017 PMID: 27119637

344. Sokolowski F, Modler AJ, Masuch R, Zirwer D, Baier M, Lutsch G, et al. Formation of critical oligomers is a key event during conformational transition of recombinant Syrian hamster prion protein. J Biol Chem. 2003; 278: 40481–40492. https://doi.org/10.1074/jbc.M304391200 PMID: 12917432

345. Bjorndahl TC, Zhou GP, Liu X, Perez-Pineiro R, Semenchenko V, Saleem F, et al. Detailed biophysical characterization of the acid-induced PrPᶜ to PrPᵝ conversion process. Biochemistry. 2011; 50: 1162–1173. https://doi.org/10.1021/bi101435c PMID: 21189021

346. Larda ST, Simonetti K, Al-Abdul-Wahid MS, Sharpe S, Prosser RS. Dynamic equilibria between monomer and oligomeric misfolded states of the mammalian prion protein measured by ¹⁹F NMR. J Am Chem Soc. 2013; 135: 10533–10541. https://doi.org/10.1021/ja404584s PMID: 23781904

347. Singh J, Udgaonkar JB. Dissection of conformational conversion events during prion amyloid fibril formation using hydrogen exchange and mass spectrometry. J Mol Biol. 2013; 425: 3510–3521. https://doi.org/10.1016/j.jmb.2013.06.009 PMID: 23811055

348. Singh J, Udgaonkar JB. Structural effects of multiple pathogenic mutations suggest a model for the initiation of misfolding of the prion protein. Angew Chem Int Ed. 2015; 54: 7529–7533.

349. Trevitt CR, Hosszu LLP, Batchelor M, Panico S, Terry C, Nicoll AJ, et al. N-terminal domain of prion protein directs its oligomeric association. J Biol Chem. 2014; 289: 25497–25508. https://doi.org/10.1074/jbc.M114.566588 PMID: 25074940

350. Groverman BR, Dolan MA, Taubner LM, Kraus A, Wicker RB, Caughey B. Parallel in-register intermolecular β-sheet architectures for prion-seeded prion protein (PrP) amyloids. J Biol Chem. 2014; 289: 24129–24142. https://doi.org/10.1074/jbc.M114.578344 PMID: 25028516

351. Nyström S, Hammarström P, Generic amyloidogenicity of mammalian prion proteins from species susceptible and resistant to prions. Sci Rep. 2015; 5: 10101. https://doi.org/10.1038/srep10101 PMID: 25960067

352. Vázquez-Fernández E, Vos MR, Afanasjev P, Cebey L, Sevillano AM, Vidal E, et al. The structural architecture of an infectious mammalian prion using electron cryomicroscopy. PLoS Pathog. 2016; 12: e1005835. https://doi.org/10.1371/journal.ppat.1005835 PMID: 27606840

353. Terry C, Wenborn A, Gross N, Sells J, Joiner S, Hosszu LLP, et al. *Ex vivo* mammalian prions are formed of paired double helical prion protein fibrils. Open Biol. 2016; 6: 160035. https://doi.org/10.1098/rsob.160035 PMID: 27249641

354. Langedijk JPM, Fuentes G, Boshuizen R, Bonvin AMJJ. Two-rung model of a left-handed β-helix for prions explains species barrier and strain variation in transmissible spongiform encephalopathies. J Mol Biol. 2006; 360: 907–920. https://doi.org/10.1016/j.jmb.2006.05.042 PMID: 16782127

355. Govaerts C, Wille H, Prusiner SB, Cohen FE. Evidence for assembly of prions with left-handed β-helices into trimers. Proc Natl Acad Sci. 2004; 101: 8342–8347. https://doi.org/10.1073/pnas.0402254101 PMID: 15155909

356. Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R, Meier BH. Amyloid fibrils of the HET-s (218–289) prion form a beta solenoid with a triangular hydrophobic core. Science. 2008; 319: 1523–1526. https://doi.org/10.1126/science.1151839 PMID: 18339938

357. Van Melckebeke H, Wasmer C, Lange A, Elso AB, Loquet A, Böckmann A, et al. Atomic-resolution three-dimensional structure of HET-s(218–289) amyloid fibrils by solid-state NMR spectroscopy. J Am Chem Soc. 2010; 132: 13765–13775. https://doi.org/10.1021/ja104213j PMID: 20828131

358. Schütz AK, Soragni A, Hornemann S, Aguzzi A, Ernst M, Böckmann A, et al. The amyloid-Congo red interface at atomic resolution. Angew Chem Int Ed. 2011; 50: 5056–5960.

359. Sabaté R, Castillo V, Espargaró A, Saupe SJ, Ventura S. Energy barriers for HET-s prion forming domain amyloid formation. FEBS J. 2009; 276: 5053–5064. https://doi.org/10.1111/j.1742-4658.2009.07202.x PMID: 19682303

360. Gendoo DMA, Harrison PM, Origins and evolution of the HET-s prion-forming protein: searching for other amyloid-forming solenoids. PLoS ONE. 2011; 6: e27342. https://doi.org/10.1371/journal.pone.0027342 PMID: 22096554

361. Daskalov A, Gantner M, Aulikki Wälti M, Schmidlin T, Chi CN, Wasmer C, et al. Contribution of specific residues of the β-solenoid fold to HET-s prion function, amyloid structure and stability. PLoS ONE. 2014; 10: e1004158.

362. Wasmer C, Zimmer A, Sabaté R, Soragni A, Saupe SJ, Ritter C, et al. Structural similarity between the prion domain of HET-s and a homologue can explain amyloid cross-seeding in spite of limited sequence identity. J Mol Biol. 2010; 402: 311–325. https://doi.org/10.1016/j.jmb.2010.06.053 PMID: 20600104

363. Mizuno N, Baxa U, Steven AC. Structural dependence of HET-s amyloid fibril infectivity assessed by cryoelectron microscopy. Proc Natl Acad Sci USA. 2011; 108: 3252–3257. https://doi.org/10.1073/pnas.1011342108 PMID: 21300906

364. Wasmer C, Soragni A, Sabaté R, Lange A, Riek R, Meier BH. Infectious and noninfectious amyloids of the HET-s(218–289) prion have different NMR spectra. Angew Chem Int Ed. 2008; 47: 5839–5841.

365. Wan W, Wille H, Stöhr J, Baxa U, Prusiner SB, Stubbs G. Degradation of fungal protein HET-s(218–289) induces formation of a generic amyloid fold. Biophys J. 2012; 102: 2339–2344. https://doi.org/10.1016/j.bpj.2012.04.011 PMID: 22677387

366. Wan W, Stubbs G. Fungal prion HET-s as a model for structural complexity and self-propagation in prions. Proc. Natl. Acad. Sci. USA. 2014; 111: 5201–5206. https://doi.org/10.1073/pnas.1322933111 PMID: 24706820

367. Wan W, Bian W, McDonald M, Kijac A, Wemme DE, Stubbs G. Heterogeneous seeding of a prion structure by a generic amyloid form of the fungal prion-forming domain HET-s(218–289). J Biol Chem. 2013; 288: 29604–29612. https://doi.org/10.1074/jbc.M113.505511 PMID: 23986444

368. Berthelot K, Immel F, Géan J, Lecomte S, Oda R, Kauffmann B, et al. Driving amyloid toxicity in a yeast model by structural changes: a molecular model. FASEB J. 2009; 23: 2254–2263. https://doi.org/10.1096/fj.08-125724 PMID: 19228879

369. Berthelot K, Ta HP, Géan J, Lecomte S, Cullin C. *In vivo* and *in vitro* analyses of toxic mutants of HET-s: FTIR antiparallel signature correlates with amyloid toxicity. J Mol Biol. 2011; 412: 137–152. https://doi.org/10.1016/j.jmb.2011.07.009 PMID: 21782829

370. Berthelot K, Cullin C, Lecomte S. What does make an amyloid toxic: morphology, structure or interaction with membrane? Biochimie. 2013; 95: 12–19. https://doi.org/10.1016/j.biochi.2012.07.011 PMID: 22824150