CrossMark

A β-solenoid model of the Pmel17 repeat domain: insights to the formation of functional amyloid fibrils

Nikolaos N. Louros[1] · Fotis A. Baltoumas[1] · Stavros J. Hamodrakas[1] ·
Vassiliki A. Iconomidou[1]

Abstract Pmel17 is a multidomain protein involved in biosynthesis of melanin. This process is facilitated by the formation of Pmel17 amyloid fibrils that serve as a scaffold, important for pigment deposition in melanosomes. A specific luminal domain of human Pmel17, containing 10 tandem imperfect repeats, designated as repeat domain (RPT), forms amyloid fibrils in a pH-controlled mechanism in vitro and has been proposed to be essential for the formation of the fibrillar matrix. Currently, no three-dimensional structure has been resolved for the RPT domain of Pmel17. Here, we examine the structure of the RPT domain by performing sequence threading. The resulting model was subjected to energy minimization and validated through extensive molecular dynamics simulations. Structural analysis indicated that the RPT model exhibits several distinct properties of β-solenoid structures, which have been proposed to be polymerizing components of amyloid fibrils. The derived model is stabilized by an extensive network of hydrogen bonds generated by stacking of highly conserved polar residues of the RPT domain. Furthermore, the key role of invariant glutamate residues is proposed, supporting a pH-dependent mechanism for RPT domain assembly. Conclusively, our work attempts to provide structural insights into the RPT domain structure and to elucidate its contribution to Pmel17 amyloid fibril formation.

Introduction

Pmel17 (known also as PMEL, Silver, SILV, gp100, and ME20) is a multidomain type I transmembrane glycoprotein, expressed in lysosome-like organelles of melanocyte cells, known as melanosomes [1–3]. Pmel17 self-assembles into intraluminal filament arrays [3, 4] which are widely considered as functional protective amyloid [5–7], an emerging concept dictating that organisms often exploit the intriguing structural properties of amyloids in order to support important biological processes [8–10]. This fibrillar matrix has an important role in pigment biosynthesis, since it serves as a template for deposition of melanin and possibly contributes in sequestration of toxic intermediates produced during melanin synthesis [5]. After expression, Pmel17 is heavily glycosylated within the ER and Golgi apparatus [11]. Subsequently, it is cleaved by a proprotein convertase (PC) into two individual fragments, an N-terminal luminal fragment and a C-terminal membrane-bound part, designated as M_α and M_β, respectively (Fig. 1) [12, 13]. The former is the polymerizing component leading to Pmel17 fibril formation and is composed of three individual domains, namely the N-terminal region (NTR), a polycystic kidney disease-like domain (PKD) and the repeat domain (RPT) [14, 15]. The M_β fragment consists of a Kringle-like domain (KRG) [16] and a C-terminal fragment, containing the transmembrane and cytoplasmic domain (TM and CTF, respectively) [15]. The PKD

✉ Vassiliki A. Iconomidou
veconom@biol.uoa.gr

1   Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Panepistimiopolis, 157 01 Athens, Greece

domain has been postulated to promote intracellular trafficking of Pmel17 to intralumenal vescicles [17]. Additionally, studies have also proposed that it is part of the Pmel17 amyloid fibril core, with the NTR and RPT domain having a regulatory role [15, 18]. On the other hand, several experimental studies indicate that the RPT domain forms amyloid-like fibrils in vitro [19–25], whereas supporting evidence indicates that deletion of the RPT leads to complete loss of Pmel17 fibril formation [6]. Interestingly, the RPT domain, containing 10 tandem imperfect repeats (Fig. 1b), has been proposed to participate in a pH-dependent mechanism of Pmel17 amyloid fibril formation [21–23], albeit no experimental evidence suggest such an in vivo mechanism [15]. Specifically, experimental results indicate that RPT fibrils are formed in vitro only under mildly acidic conditions and completely dissolve when the pH is increased [23, 24]. This reversible mechanism has been proposed to have important biological implications by allowing Pmel17 aggregation only under the mildly acidic melanosome pH ($\sim$4.5–5.5) and concurrently protecting from cytotoxic effects by preventing aggregation under the neutral cytosolic pH [22].

The three-dimensional structure of the RPT domain of Pmel17 is currently unknown. However, several theoretical and experimental studies have indicated that proteins composed of repeats of 5–40 residues, such as the RPT domain, most commonly form closed superhelical structures, known as solenoids [26–29]. More specifically, repeats at the lower end of that range usually lead to the formation of β-solenoid structures, which have been previously associated with being polymerizing components of amyloids [29–32]. The polypeptide chain coils around the axis of the solenoid forming successively stacked coils, each composed of single or multiple repeats. Since β-solenoid repeats consist of several polar residues, the stacked coils are coordinated and organized by a network of H-bonds, formulated between the polar side chains and with the polypeptide backbone [26–29]. Analysis has previously indicated that β-solenoid proteins presenting sequence identity close to or above 18 % tend to have similar folds [33]. In our work, by performing structural and computational analysis of the RPT domain, we propose that the RPT domain of Pmel17 folds as a β-solenoid structure, stabilized by an extensive network of hydrogen-bonded ladders formed by conserved stacked polar residues. Furthermore, we attempt to elucidate how the characteristic β-solenoid properties of the RPT domain may have both functional and structural implications in the formation of the Pmel17 fibrillar matrix. Briefly, our results provide structural insights into the RPT domain signifying its importance in the formation of Pmel17 amyloid fibrils.
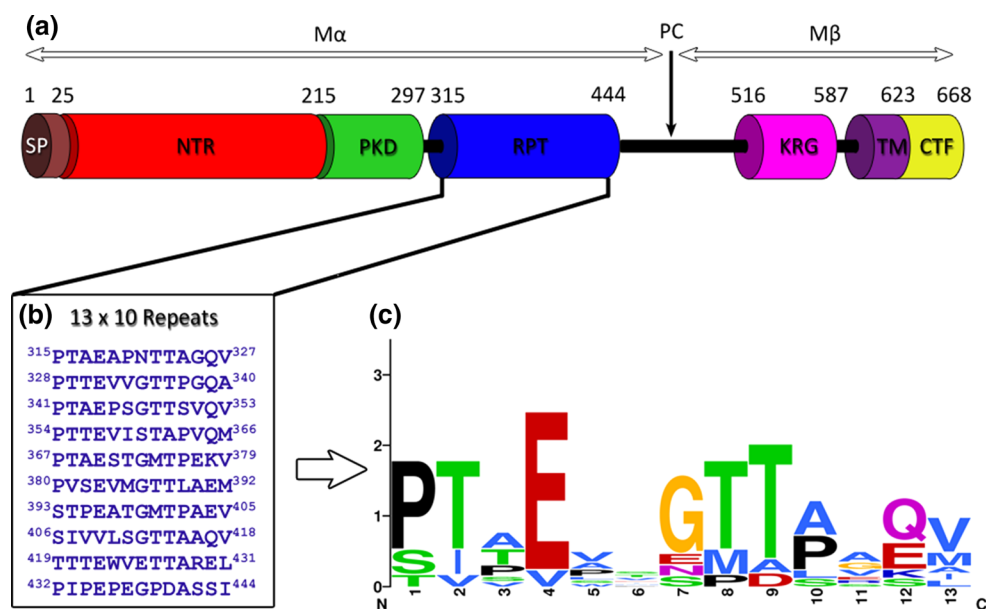

Fig. 1 Schematic representation of the Pmel17 multidomain protein. a Pmel17 is cleaved by a proprotein convertase (PC) into two major fragments, known as M_α and M_β, respectively. The former, considered responsible for the formation of Pmel17 amyloid fibrils, contains three distinct domains, the NTR domain, the PKD domain and the RPT domain. b The RPT domain is composed of ten tandem imperfect repeats, composed of 13 residues each. c The characteristic pattern of the Pmel17 tandem repeats. The strong presence of polar residues, at the expense of hydrophobic or aromatic residues is clearly seen. Furthermore, the high content of Pro residues, in addition to the highly conserved Glu residues is also highlighted

155

Materials and methods

Multiple sequence alignment

Sequence segments of the Pmel17 protein from several vertebrate species, corresponding to the RPT domain, were extracted from Uniprot [34], with the following accession numbers: P40967 (*Homo sapiens*), I0FGN0 (*Macaca mulatta*), H2Q669 (*Pan troglodytes*), L8YCF6 (*Tupaia chinensis*), A0A0D9QZ09 (*Chlorocebus sabaeus*), A0A096NUA0 (*Papio anubis*), H2NHM7 (*Pongo abelii*), G3TDI1 (*Loxodonta africana*), G1L8V7 (*Ailuropoda melanoleuca*), M3WJU4 (*Felis catus*), F7HR25 (*Callithrix jacchus*), G1P0J6 (*Myotis lucifugus*), M3XPD4 (*Mustela putorius furo*), H0WVH1 (*Otolemur garnettii*), G1SEI3 (*Oryctolagus cuniculus*), W5PM53 (*Ovis aries*), H0VFQ9 (*Cavia porcellus*), Q60696 (*Mus musculus*), Q06154 (*Bos taurus*), Q98917 (*Gallus gallus*), A7XY82 (*Canis familiaris*), F7DCG1 (*Equus caballus*), L5KV09 (*Pteropus alecto*), G5BB96 (*Heterocephalus glaber*), L8I091 (*Bos mutus*), J9RN89 (*Capra hircus*). Initial detection of sequence repeats was performed by applying TRUST [35], Internal Repeats Finder [36], HHrepID [37] and RADAR [38], tools that specialize in detecting tandem and divergent sequence repeats. Multiple sequence alignment of the species-specific repetitive segments identified was subsequently performed utilizing ClustalW [39] and Jalview [40]. Finally, all species repeat segments were fitted to the well-reviewed repetitive segments of human Pmel17 by applying ClustalW and manual editing utilizing Jalview, resulting to the final multiple sequence alignment of the RPT domains (Fig. 2).

Prediction of RPT domain β-solenoid propensity

The REPETITA algorithm was applied [41], in order to identify the emerging propensity of the human RPT domain (315–444) of Pmel17 to adopt a β-solenoid structure. Specifically, solenoid prediction was performed by applying the REPETITA web tool (http://protein.bio.unipd.it/repetita) on the corresponding sequence of the RPT domain (Fig. 3).

Comparative modelling of the RPT β-solenoid structure

RepeatsDB was scanned in order to identify a suitable β-solenoid structure that could serve as template [42]. As a result, the crystal structure (PDB ID: 1P9H) of the collagen-binding domain of *Yersinia enterocolitica* adhesin (YadA) was selected [43]. Sequence alignment of the first five consecutive repeats of Pmel17 to the β-solenoid domain of the aforementioned crystal structure of YadA was performed using ClustalW [39] and manual editing utilizing Jalview [40]. The sequence alignment was guided by previous experimental data regarding the RPT domain, derived through NMR [19]. Specifically, previous detailed experiments have indicated that all hydrophobic Val and Leu residues, in addition to most of the Met residues found within the RPT repeats of Pmel17 are in extended (β-strand) conformations. Correspondingly, fitting of the RPT to YadA repeats was performed in an attempt to meet these structural demands. The five RPT repeat modelled ensemble was utilized as a template, during a second round of modelling for the remaining five C-terminal RPT repeats. Based on the above, a three-dimensional model of the RPT domain of Pmel17 (315–444) was constructed. The final repeat (431–444) has an extremely high content in proline (Pro) and negatively charged residues (seven out of thirteen residues). This indicates that the last repetitive segment is most probably unstructured and does not contribute to the β-solenoid fold but probably acts as a C-terminal cap [29]. Threading was carried out utilizing Modeller9v12 [44].

Molecular dynamics simulations

A number of molecular dynamics simulations were performed on the produced model for the refinement and subsequent evaluation of the Pmel17 RPT domain structural features. The generated model was initially refined through a short simulated annealing protocol with harmonic restraints. The model was subjected to thorough conjugate gradient energy minimization. Subsequently the system was gradually heated from 0 to 300 K and was left to gradually cool down over 30 ps, followed by a final short energy minimization. All solvent exposed glutamate (Glu) and Asp residues were protonated and harmonic restraints were applied to secondary structure elements throughout the simulations. Refinement was performed using the Generalized Born Implicit Solvent model [45] to describe the environment and the CHARMM36 all-atom force field [46, 47].

Subsequent simulations used the Protein Atoms in Coarse-Grained Environment (PACE) force field to model the system [48]. PACE is a hybrid model, combining representations in different resolutions to enable fast and accurate simulations of biological systems and processes. A united-atom representation is used to describe the proteins with all heavy atoms and polar hydrogens modelled explicitly, while the solvent and ions are described using the MARTINI Coarse-Grained force field [49], in which an interaction particle describes, on average, four heavy atoms. PACE has been parameterized to achieve accurate reproduction of experimental thermodynamic data as well
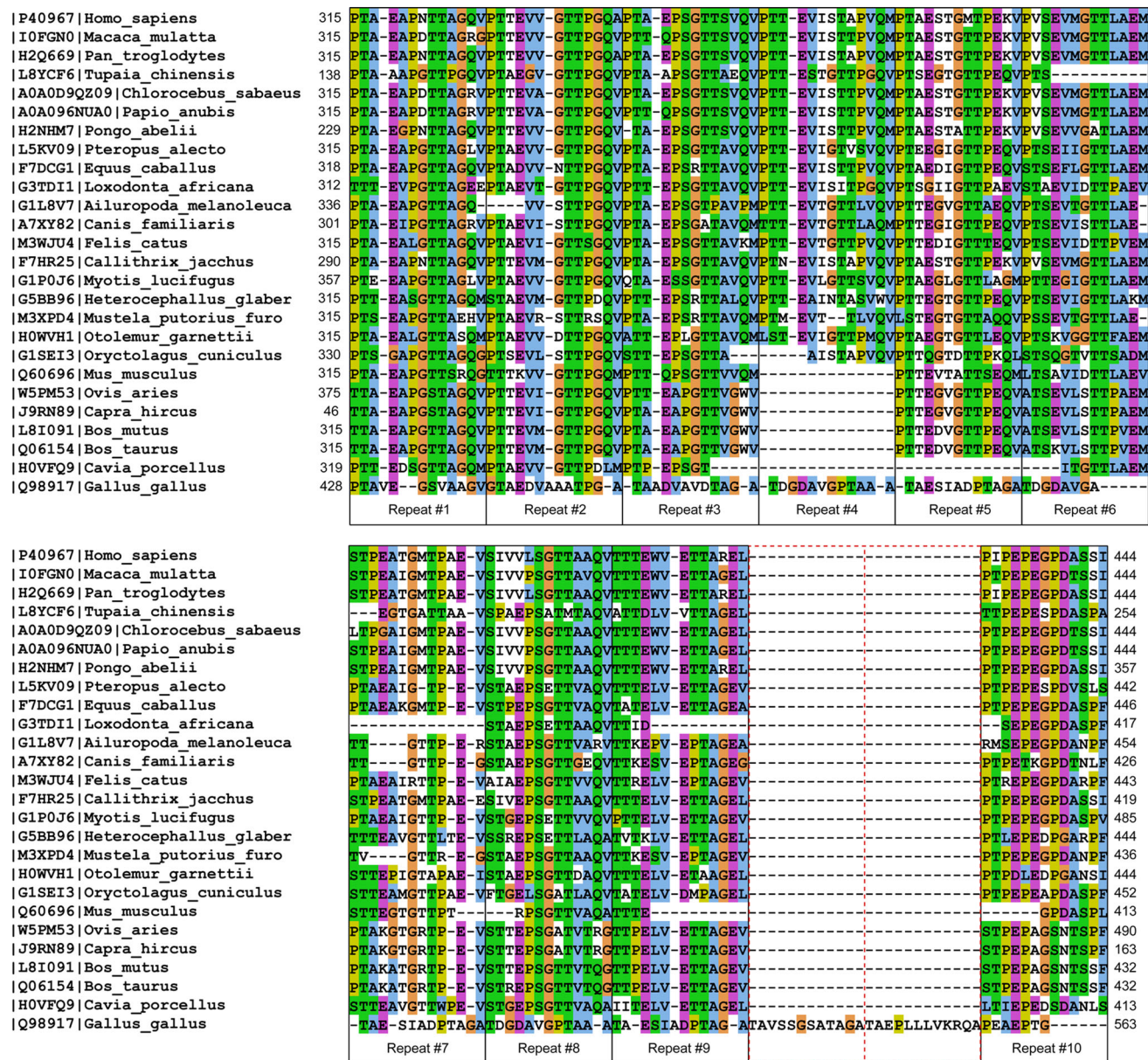
Fig. 2 Multiple sequence alignment of the RPT domain from several vertebrate species. RPT domain regions are defined by the sequence identifiers right before and after each aligned sequence. Subsequent tandem repetitive units are enclosed in *black boxes* and numbered sequentially. All species contain an RPT domain composed of ten non-perfect 13-residue long repeats, except of *O. aries, B. Taurus, B. mutus, C. hircus* and *M. musculus* that are missing the fourth repetitive unit, *L. africana* missing the seventh and *C. porcellus* missing both the fourth and fifth unit, in addition to half of the third and sixth repeat, respectively. The evolutionary divergent chicken RPT domain is composed of 12-residue repeats. Additionally, it contains two additional repeats, shown in *dotted red boxes*, interposed between the ninth and final repeat of the other vertebrate RPT domains. The repeat units consist of highly conserved polar or negatively charged residues and exhibit a strong preference for Gly and Pro residues, respectively. Finally, the C-terminal repeat unit presents significantly lower conservation, however displays an increased content in Pro or negatively charged residues in all species

as results from all-atom simulations, while at the same time enabling an about 30-times increase in simulation performance. The robustness of the force field has been validated for a diverse range of proteins and processes, including the equilibration of α-helical and β-barrel transmembrane proteins embedded in lipid bilayers and micelles [50], the successful folding of small all-α and all-β protein domains

[48, 51] and the thermodynamics of fibril elongation in amyloid-β aggregates [52].

The lowest energy model obtained from refinement was used to build the initial systems. The PACE v1.3 united atom model was used to describe the protein, while MARTINI v. 2.0 was used to model the solvent [48, 49]. The model was inserted in a $100 \times 100 \times 120$ Å water
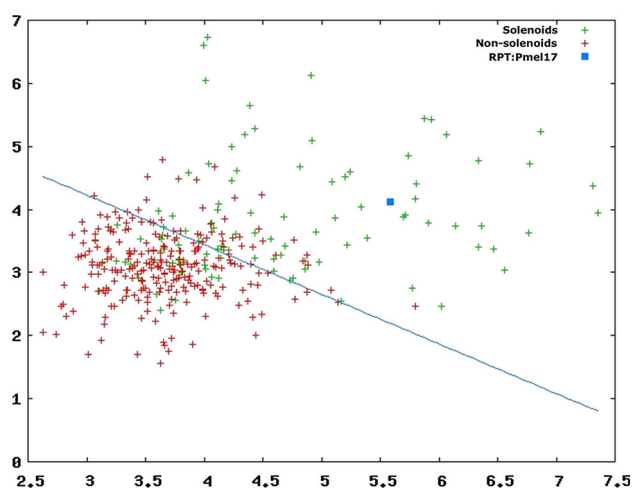
Fig. 3 Output of the REPETITA method [41]. The scatter plot y-axis contains information of the optimal θ-ratio versus Zmax (x-axis) for the training sets and the RPT domain of human Pmel17 (residues 315–444). *Red crosses* correspond to non-solenoid proteins of the training set, whereas *green crosses* correspond to solenoid structures, respectively. The RPT domain prediction, indicated with a *blue square*, clearly falls within the region of solenoid structures

box and all solvent exposed Glu and Asp residues were protonated. The system was subjected to 5000 steps of conjugate gradient energy minimization and was subsequently equilibrated over 1 ns, applying harmonic constraints of gradually decreasing strength on the protein atoms. Finally, the equilibrated systems were simulated without any restraints for 100 ns. All simulations were performed in the isothermal–isobaric ensemble (NPT) with temperature and pressure controlled at 310 K and 1.013 bar (∼1 atm.), respectively, using periodic boundary conditions and a 5 fs timestep.

Three independent simulations, initiated with randomized velocity, were carried out for the system using the described simulation protocol. The last 10 ns from each simulation result were concatenated and clustered using the quality threshold clustering algorithm [53]. The root mean square deviation (RMSD) between frames was used as the distance function during clustering, with a cut-off value of 1.0 Å between superimposed atoms. The centroids of the clusters representing the top 95 % of the analyzed trajectories were extracted and the lowest energy conformation was chosen to represent the final model.

An important limitation of the generalized born solvent, as well as the hybrid united-atom/Coarse-Grained representation adopted by the PACE simulations is the modelling of electrostatic interactions. Non-bonded interactions are calculated using finite cut-offs and shifted potentials, and, in particular, Coulomb electrostatics are screened implicitly. While this enables faster simulations, it precludes the accurate estimation of the electrostatic potential,

particularly for long-range interactions, which have been found to have significant impact upon many biological processes. As such, the effects of low pH conditions and Glu protonation upon the Pmel17 RPT domain may not be modelled adequately in a hybrid resolution model.

To further evaluate the contribution of electrostatics and the importance of Glu protonation, the stability of the RPT domain was evaluated with all-atom molecular dynamics simulations in explicit solvent. The model generated by clustering the PACE simulations was used as the basis for the simulations. Two different systems were designed, featuring either protonated or unprotonated solvent exposed Glu and Asp residues. The CHARMM36 all-atom force field and the TIP3P water model were used in the simulations [46, 47]. Each model was embedded in a 100 × 100 × 120 Å solvent box, built with the TIP3P water model and NaCl counterions. The systems were energy minimized for 5000 steps and were subsequently equilibrated over 1 ns, applying gradually decreasing restraints on the proteins. Finally, each system was simulated without restraints. Simulations were performed in the NPT ensemble, with temperature and pressure maintained at 310 K and 1 atm, respectively. Periodic boundary conditions were used and the Particle Mesh Ewald method was applied to screen long range electrostatic interactions. Two independent simulations were run for each system using the described protocol, initialized with randomized velocity.

All molecular dynamics simulations were performed with NAMD Scalable Dynamics [54]. Simulation results were analyzed using visual molecular dynamics (VMD) [55], carma trajectory analysis [56] and in-house developed scripts. The final models were further validated using DSSP [57] and the WHAT-IF package [58]. Structural alignments were performed using Dali [59, 60] and VMD. Visualization and rendering were performed using VMD and PyMOL [61]. Detailed information on the simulations can be found in the Supplementary Material, available online.

Results and discussion

Evolution of RPT domain repeat ensemble provides evidence of a β-solenoid fold

Multiple sequence alignment of several vertebrate RPT domains revealed that it is a variable region, in accordance to previous studies suggesting that protein repeat segments erode fast [62]. Significant differences emerged regarding the number of repeats that are present in-between species. Specifically, most species have a fixed number of ten repeats composing the RPT domain, with the exception of

six mammalian species, namely *O. aries*, *B. taurus*, *B. mutus*, *C. hircus*, *M. musculus* and *L. africana,* which are missing a single repetitive segment and *C. porcellus* which is missing three repetitive segments. Additionally, the evolutionary distant chicken RPT domain (*G. gallus*) presents two additional repeats and is, therefore, composed of 12 repetitive segments, respectively (Fig. 2). Impressively, domains presenting incremental additions/deletions of repeats, such as the RPT domain, usually occur in linear structural arrangements, such as closed superhelical structures, known as solenoids that can accommodate the loss or gain of repetitive segments by prolonging or limiting their linear length [27, 62]. Supporting information is derived by multiply aligning the subsequent cross-species sequence repeats. Analysis indicated that the repeat length is conserved between species to an average of 13 residues, with the exception of the chicken RPT domain which is composed primarily of 12-residue repeats. Such repeat lengths fall within the length region (5–40 residues) of repetitive segments commonly associated with the formation of β-solenoids [26–29].

Valuable information is extracted through analysis of the amino acid composition and conservation of the RPT repeats. The amino acid composition between species revealed an extremely high content of residues with polar side chains, in addition to the strong presence of glycine (Gly) and Pro residues, respectively. Importantly, this correlates to the amino acid composition of β-solenoids. Such structures usually present a preference in residues with polar side chains that are buried within the hydrophobic core by formulating hydrogen bond ladders, as well as a high content in Gly residues [26–29]. Moreover, impressive conservation is also witnessed for the Glu residues of each repetitive segment. These residues have been previously proposed to have an important role in the formation of RPT amyloid fibrils [22, 25]. Furthermore, Pro residues are also highly conserved, although they are uncommon in amyloid-forming proteins (AFPs), since they act as β-breakers and do not facilitate the "cross-β" fold of amyloids. However, they are frequently accommodated in β-solenoid structures by stacking interactions formed along the β-helical axis [63]. Conclusively, the RPT domain of Pmel17 shares several common properties to β-solenoid structures.

Impressively, theoretical studies have previously postulated that β-solenoid structures could actually be the polymerizing components of amyloid fibrils [8, 29–32]. Structural studies have associated several AFPs with a β-solenoid conformation or with β-solenoid-like structures formed by stacked β-hairpins [8, 64–66]. Additionally, structures of AFPs, obtained by solid state NMR promote this suggestion, such as in the case of HET-s or Aβ$_{1-40}$ [67, 68]. Finally, a β-solenoid fold for the human RPT domain

is further supported by utilizing the REPETITA algorithm which specializes in discriminating β-solenoid from non-solenoid protein sequences. The REPETITA prediction, regarding the human RPT domain ($\rho_\theta = 4.127$ and $Z_{max} = 5.584$) undoubtedly falls within the region of β-solenoid protein sequences, suggesting a β-solenoid fold with high certainty for the RPT domain, based on the distance (1.52) of the prediction from the optimal line separating β-solenoid from non-solenoid sequences (Fig. 3).

Modelling of the human RPT domain β-solenoid structure

A three dimensional model of the human RPT domain was constructed utilizing the crystal structure of the collagen binding domain of YadA as a template [43]. This selection was based primarily on the similar length of the repetitive segments (13–14 aa) YadA comprises and the relative homology it presents to the RPT domain. Additionally, our selection was also supported by the fact that the template structure is composed of short β-strands, since the high Pro content of the RPT domain restricts the formation of long uninterrupted β-strands (Fig. 2). The derived model indicates that the human RPT domain folds into a left-handed β-solenoid structure, which consists of 10 coils, formed by each successive repetitive unit (Fig. 4b, c). The C-terminal repeat of Pmel17, containing four Pro residues and three negatively charged residues is, however, unstructured. Supporting studies indicate that most β-solenoids are C-terminally capped by terminal repeats commonly containing a high number of Pro residues, in order to forestall polymerization [29, 69, 70]. Structural alignment experiments validate our model by indicating that it closely resembles the template structure (RMSD 1.0 Å).

The stability and dynamics of the model were validated by explicit solvent molecular dynamics simulations under constant temperature and pressure conditions. Simulations were performed using the PACE hybrid resolution force field, combining a united-atom representation of protein atoms in near atomic detail with the use of a Coarse-Grained solvent. Three independent 100 ns simulations were carried out for the system and the results were evaluated by monitoring time-dependent geometrical properties (Fig. S1-S4). The final configurations from each simulation display small deviations from each other. The model appears to reach equilibrium within the first 20 ns in each simulation, with only minor changes observed with respect to the initial protein coordinates, as indicated by the minor (1.5–2.0 Å) changes in RMSD measured with respect to the starting conformation (Fig. S1). Calculation of RMSD with respect to the average structure, as well as the radius of gyration show minor deviations as well after the first
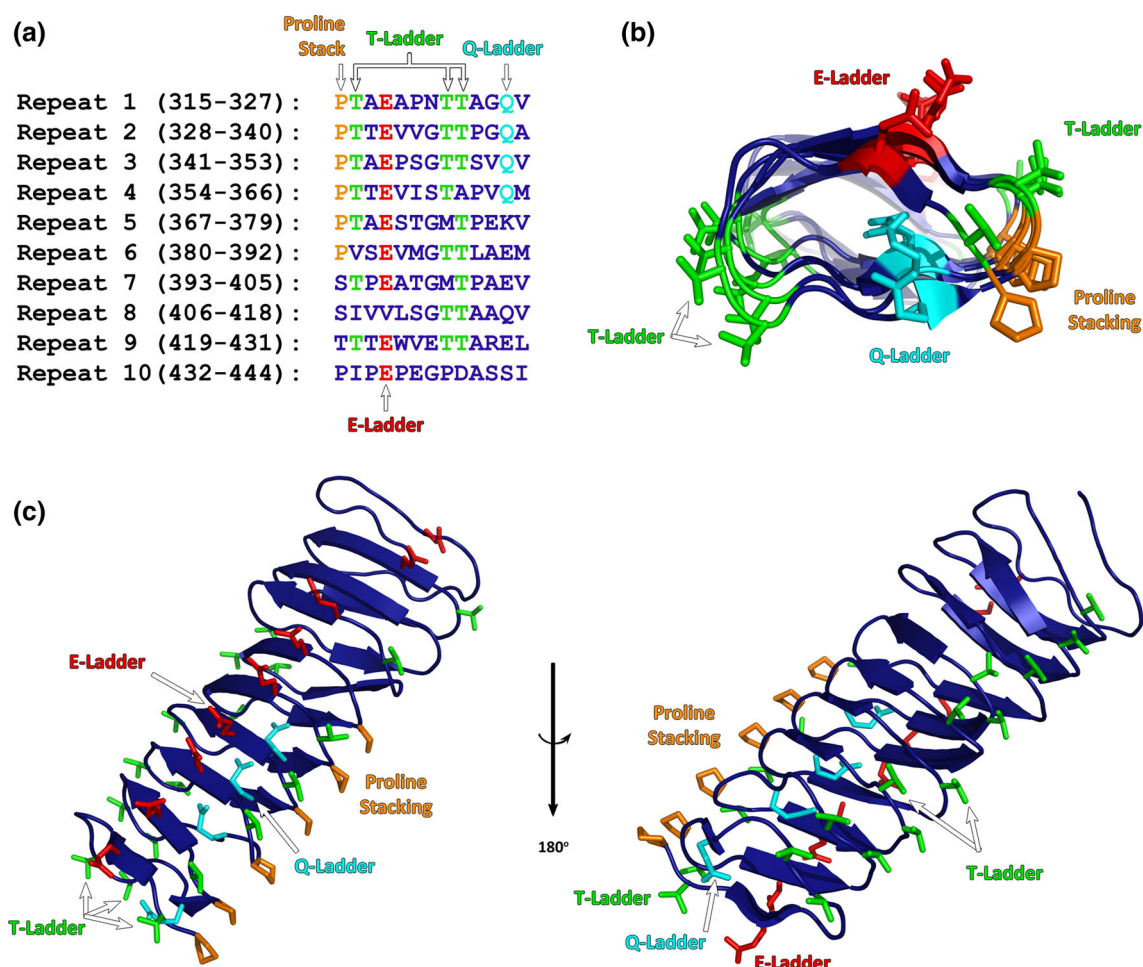
Fig. 4 Properties of the RPT β-solenoid structure, derived from extensive coarse-grain MD simulations. **a** Positions of the highly conserved Pro, Thr, Glu and Gln residues, playing an important role in the formation of the β-solenoid structure, are highlighted in *orange*, *green*, *red* and *cyan colour*, respectively. **b** View along the axis of the β-solenoid structure indicates that it is an O-type β-solenoid structure, based on the coil shape. Positions of the aforementioned conserved residues are also indicated in *sticks* and *colour-coded* as in (**a**). **c** View along the β-solenoid axis of the final RPT β-solenoid structure. Hydrogen bond ladders are formed along the solenoid axis by the conserved Thr (T-ladders), Glu (E-ladder) and Gln (Q-ladder) residues, whereas the conserved Pro residues are stacked at the exterior, stabilizing the structure and concurrently forming a potential interaction interface

20 ns, suggesting the optimal packing quality of the model (Fig. S2-S3). Root mean square fluctuation (RMSF) analysis shows that the most profound fluctuations were observed mainly for the N- and C-terminal regions of the domain (Fig. S4), while the overall β-solenoid fold and secondary structure elements remained intact.

Backbone hydrogen bonds between β-strands were retained, with small fluctuations in average length of each strand, as shown by secondary structure analysis (Fig. S5). An exception to this case is the 394–410 region of the model, in which the β-structure is slightly distorted. This is caused by the presence of a Proline (Pro-395), which precludes the formation of hydrogen bonds between the two β-strands in the sheet. This results in a slightly shorter β-strand for the 395–400 region and a slight twist between the 315–395 part and the rest of the model. However, the

structure seems to be stabilized by a salt bridge, formed by the buried Glu-404 with a similarly buried Lysine (Lys-378) and, as a result, the overall β-solenoid structure remains stable. The lowered pH conditions would constitute the formation of a salt bridge between the two proposed residues improbable. However, both residues are buried within the core of the β-solenoid fold. Therefore, it is possible that the specific salt bridge is formed once the general fold has occurred, within the hydrophobic core, where the individual residues are no longer affected by the acidic solvent conditions. Based on the above, it seems that the salt bridge has a more stabilizing role in the final fold of the RPT domain rather than guiding its formation. It should be noted that this theoretical model meets most of the restraints derived from solid-state NMR experiments [23, 24]; Met, Leu and Val residues are almost always in an

extended conformation, with only one Val residue (Val-381) being the exception, appearing in a turn between two β-strands. This conformation was largely retained throughout the 100 ns simulations (Fig. S5), with occasional fluctuations observed for residues in the β-strand ends.

Interestingly, the close packing of Pro, Glu, Gln and Thr residues and the associated ring stacking and hydrogen bond ladders, originally observed in the starting model, were also retained throughout the simulations. Close packing interactions formed by these residues appear to be pertinent to the RPT domain solenoid structure and are presented in detail in the following sections. Overall, these observations validate the stability and quality of the proposed β-solenoid structure model for the RPT domain.

Role of polar stacking in the RPT β-solenoid structure

The derived β-solenoid structural model of the RPT domain shares several common and distinguishable features of the β-solenoid fold. It is a left-handed O-type β-solenoid structure, with an extremely tight hydrophobic core due to the small size of the repetitive coils. The core is shaped mostly by repeat positions of aliphatic residues and stacked glutamine residues forming an inward H-bond ladder, located in the centre of the structure (Fig. 4). Burial of their polar side chains is compensated by the formation of H-bond ladders, a striking feature that further stabilizes β-solenoid structures [28, 29]. Additionally, the structure is also stabilized by the formation of three H-bond ladders, located in the exterior, which are formed by stacked Thr residues (T-ladders). Apart from further stabilizing the β-solenoid fold, the T-ladders facing outwards, composed of highly invariant Thr residues (Fig. 2), most probably have an additional role, since they may possibly serve as potential O-glycosylation sites. Previous studies have shown that the RPT domain of Pmel17 is heavily O-glycosylated [11]. This process has been inferred to be critical for the ability of the protein to form amyloid fibrils [71]. Impressively, previous studies have shown that Thr stacking has also important functional implications in other β-solenoid AFPs, such as antifreeze proteins [72]. Consequently, it appears that the Thr residues, which occupy highly conserved repeat positions (Fig. 2), retain both a structural and functional role in the formation of RPT amyloid fibrils.

The β-solenoid fold of the RPT domain supports the presence of Pro residues

Excess of Pro residues has been extensively shown to impart the unfolded state of proteins [26]. Although Pro residues frequently appear in protein repeats, studies propose that they can be accommodated in β-solenoid structures. This is accomplished by stacking of Pro residues, resulting in the formation of favourable interactions between their side chains [63]. Additionally, the structural rigidity provided by the extensive H-bond network of β-solenoids may also compensate for the presence of several Pro residues [26, 27, 29]. Pro residues have been previously proposed to serve in the containment of β-sheet elongation, as in the case of human prion protein, another AFP proposed to adopt a β-solenoid fold [73]. A similar function is also seen in the RPT model. Specifically, both the conserved Pro stack and individual Pro residues of the RPT ensemble are protruding primarily at the exterior sides of the β-solenoid structure, facilitating the formation of the β-solenoid turns and serving in the containment of β-strand extensions. The final repetitive segment, rich in Pro residues, most probably serves as a β-solenoid cap, whereas the previous three C-terminal Pro-free units, are unconstrained and clearly favour a β-strand conformation (Figs. 4, 5). Pro stacking has also been extensively associated with an important role in protein–protein interactions [74, 75]. Therefore, it is possible that these residues may contribute in interactions between RPT monomers. Finally, stacked Pro residues may also have an important role in facilitating the O-glycosylation process of Pmel17, as it has been previously suggested [76].

Stacking of Glu side chains is essential for the pH-controlled formation of RPT amyloid fibrils

Detailed studies have previously postulated that the RPT domain of Pmel17 forms amyloid fibrils only in acidic environments (pH lower than 6) [23–25]. This has been suggested to be part of an elaborate mechanism that protects Pmel17 from incorporating into fibrils, until it is found in the acidic melanosome matrix (pH ∼ 4.5–5.5). Consequently, the highly invariant Glu residues of the RPT domain (Fig. 2) have been proposed to vitally contribute in this process [22, 25]. This notion is based on the fact that the pKa values of Glu side chains in proteins correspond to the 3.5–5.0 range. As a result, the Glu side chains are most likely neutralized when found within the melanosome matrix, thus promoting Pmel17 incorporation into amyloid fibrils. Evidently, this is clearly supported by the modelled structure proposed in our study. The Glu side chains stack along the axis of the β-solenoid fold (Fig. 5). Presumably, this could not be facilitated at neutral pH, since the Glu side chains in close proximity would be repulsed. However, under the acidic pH of the melanosome matrix, the Glu side chains are neutralized. Moreover, since these side chains are likely protonated [22], they most likely contribute in the formation of an exterior H-bond ladder (E-ladder) that
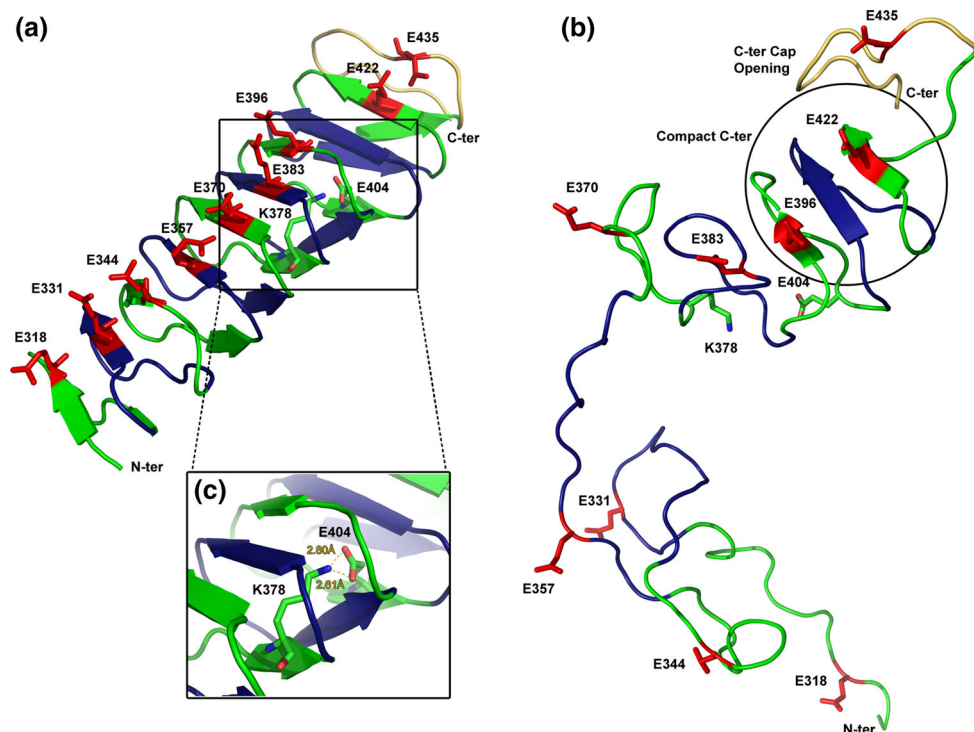
Fig. 5 An elegant pH-controlled mechanism for the β-solenoid RPT assembly, verified by all-atom MD simulations. **a** The β-solenoid is formed under acidic environmental conditions, in which the protruding conserved Glu residues of the RPT have protonated side chains. The protonated and successively stacked Glu residues possibly formulate an extra H-bond ladder in the molecular exterior side, thus providing additional stability and rigidity to the β-solenoid fold. **b** Under neutral conditions, the preformed β-solenoid fold is disrupted, due to the strong repulsive interactions between the negatively charged side chains of the otherwise stacked Glu residues, in line with detailed experimental results suggesting that RPT preformed fibrils dissolve when relocated in neutral pH conditions.

Impressively, a major part of the RPT C-terminal segment remains compact and retains the average initial fold, indicating the increased stability and possible aggregation propensity of the region, in agreement to detailed solid state NMR data highlighting this region as the major part of the immobilized core of RPT amyloid fibrils. **c** A salt bond is formed between residues K378 and E404 (shown in *sticks*) and is retained through the MD timeframe, providing stability in the overall β-solenoid fold, under acidic conditions. The RPT repeat segments are colored successively in *green* and *blue*, whereas the unstructured C-terminal cap is shown in *yellow*. The conserved Glu residues that play a major part in the pH-mediated mechanism are shown in *sticks* and highlighted in *red*

further stabilizes the β-solenoid structure (Fig. 5). In contrast, when found again under neutral conditions, the repelling interactions of the Glu side chains should lead to destabilization of the β-solenoid structure and therefore in turn of the Pmel17 amyloid fibrils. This assumption is validated by experimental evidence indicating that preformed RPT amyloid fibrils are dissolved when relocated in neutral pH [23, 24].

In order to evaluate the proposed solenoid structure against these observations, all-atom Molecular Dynamics simulations were performed in low pH conditions, in which all solvent exposed Asp and Glu residues in the model were protonated, as well as in neutral pH, where the same residues were negatively charged. Simulation results propose that side chain protonation in the Glu ladders helps the RPT solenoid retain its stability. Throughout the 40 ns all-atom simulation the model remains relatively unchanged, with fluctuations observed mainly in its N- and C-termini and an average RMSD value of 1.7 Å from the starting

conformation (Fig. S6). The solvent exposed, protonated Glu side chains remain closely packed, favouring the formation of potential hydrogen bond ladders. The intramolecular salt bond, formed between the Lys-378 and Glu-404 side chains (considered as buried residues during these simulations) is also retained throughout the simulation, indicating its important role in the stabilization of the general β-solenoid fold. Proline stacking, as well as Gln and Thr ladders are also retained. In contrast, simulations with unprotonated Glu residues led to unfolding of the solenoid within the first 20 ns, due to repulsive forces between the charged side chains in the ladders. It should be noted that unfolding is mainly observed for the 315–405 part of the solenoid that features stacking of Glu residues in ladders. In contrast, the C-terminal 405–425 region retains several of its β-solenoid properties (Figs. 5 and S7). Interestingly, solid state NMR spectra from different Pmel17 fibril populations show that this region contributes the most to the immobilized fibril core [25]. The striking

stability and compactness of this region may be a hint towards its importance in the formation of Pmel17. Moreover, the Lys-378–Glu-404 salt bond is no longer retained. The fact that Glu-404 is located right next to the border between the 315–405 and 405–425 regions of the RPT domain might suggest that this interaction contributes in stabilizing the β-solenoid fold of the less stable N-terminal region without, however, being sufficient to keep the specific fold intact under the neutral pH conditions. In any case, the observed disruptions in the solenoid structure, proposed by the simulations, would be expected to lead to disruptions in Pmel17 RPT aggregates, in line with experimental observations.

A C-terminal capping mechanism possibly mediates RPT amyloid fibril formation

Several β-solenoid domains have been associated with polymerization of AFPs into amyloid fibrils. This process is usually facilitated by head-to-tail interactions between exposed terminal coils, which eventually lead to the formation of fibrils with indeterminate length [29, 69]. Accumulating evidence reveals that polymerization is interrupted, at most cases, by N- and C-terminal solenoid capping [29, 69]. Frequently, this is facilitated by terminal coils, introducing a high content in Pro or charged residues, ending the H-bonding complementarity of β-solenoids [69, 70]. Interestingly, a similar mechanism is revealed for the RPT domain. Residues close to the C-terminal part of the RPT domain, contribute to the immobilized fibril core of RPT amyloid fibrils [25], however, this process may be mediated by the presence of the C-terminal repeat. As stated previously, this segment exhibits characteristics of a capping motif, since it is composed of several Pro residues and negatively charged residues, a feature that is more or less conserved between species (Fig. 2). Conclusively, this segment may play an important role in mediating the formation of Pmel17 amyloid fibrils by preventing polymerization of the aggregation-prone C-terminal part of the RPT domain, prior to cleavage of the Mα fragment, an essential step in Pmel17 amyloid fibril formation [71].

Conclusions

Little is known regarding the three-dimensional structure of the RPT domain of Pmel17. However, studies have extensively associated this domain with the formation of Pmel17 amyloid fibrils [19–25]. Our study provided detailed evidence indicating that this repeat segment most probably folds into a β-solenoid structure presenting all the basic distinct properties of β-solenoids. The structural rigidity and the wide, elongated surfaces of β-solenoids are important properties of biological adhesion surfaces. Therefore, a β-solenoid fold would facilitate the formation of the Pmel17 substrate, which is utterly important in biosynthesis of melanin. In summary, our findings attempt to provide structural information on the fold of the RPT domain that may elucidate its important role in the elegant and elaborate mechanism of Pmel17 amyloid fibril formation. Unveiling this reversible mechanism of functional amyloid fibril formation is of great importance, since it could shed some light in the underlined mechanisms behind functional and pathological amyloid fibril formation and have various potential applications in the field of nanobiotechnology or biomaterials.

References

1. Kushimoto T, Basrur V, Valencia J, Matsunaga J, Vieira WD, Ferrans VJ et al (2001) A model for melanosome biogenesis based on the purification and analysis of early melanosomes. Proc Natl Acad Sci USA 98(19):10698–10703. doi:10.1073/pnas.191184798

2. Raposo G, Tenza D, Murphy DM, Berson JF, Marks MS (2001) Distinct protein sorting and localization to premelanosomes, melanosomes, and lysosomes in pigmented melanocytic cells. J Cell Biol 152(4):809–824

3. Berson JF, Harper DC, Tenza D, Raposo G, Marks MS (2001) Pmel17 initiates premelanosome morphogenesis within multivesicular bodies. Mol Biol Cell 12(11):3451–3464

4. Theos AC, Tenza D, Martina JA, Hurbain I, Peden AA, Sviderskaya EV et al (2005) Functions of adaptor protein (AP)-3 and AP-1 in tyrosinase sorting from endosomes to melanosomes. Mol Biol Cell 16(11):5356–5372. doi:10.1091/mbc.E05-07-0626

5. Fowler DM, Koulov AV, Alory-Jost C, Marks MS, Balch WE, Kelly JW (2006) Functional amyloid formation within mammalian tissue. PLoS Biol 4(1):e6. doi:10.1371/journal.pbio.0040006

6. Hoashi T, Muller J, Vieira WD, Rouzaud F, Kikuchi K, Tamaki K et al (2006) The repeat domain of the melanosomal matrix protein PMEL17/GP100 is required for the formation of organellar fibers. J Biol Chem 281(30):21198–21208. doi:10.1074/jbc.M601643200

7. Hurbain I, Geerts WJ, Boudier T, Marco S, Verkleij AJ, Marks MS et al (2008) Electron tomography of early melanosomes: implications for melanogenesis and the generation of fibrillar amyloid sheets. Proc Natl Acad Sci USA 105(50):19726–19731. doi:10.1073/pnas.0803488105

8. Iconomidou VA, Vriend G, Hamodrakas SJ (2000) Amyloids protect the silkmoth oocyte and embryo. FEBS Lett 479(3):141–145

9. Iconomidou VA, Hamodrakas SJ (2008) Natural protective amyloids. Curr Protein Pept Sci 9(3):291–309

10. Shewmaker F, McGlinchey RP, Wickner RB (2011) Structural insights into functional and pathological amyloid. J Biol Chem 286(19):16533–16540. doi:10.1074/jbc.R111.227108

11. Valencia JC, Rouzaud F, Julien S, Chen KG, Passeron T, Yamaguchi Y et al (2007) Sialylated core 1 O-glycans influence the sorting of Pmel17/gp100 and determine its capacity to form fibrils. J Biol Chem 282(15):11266–11280. doi:10.1074/jbc.M608449200

12. Berson JF, Theos AC, Harper DC, Tenza D, Raposo G, Marks MS (2003) Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis. J Cell Biol 161(3):521–533. doi:10.1083/jcb.200302072

13. Leonhardt RM, Vigneron N, Rahner C, Cresswell P (2011) Proprotein convertases process Pmel17 during secretion. J Biol Chem 286(11):9321–9337. doi:10.1074/jbc.M110.168088

14. Theos AC, Truschel ST, Raposo G, Marks MS (2005) The Silver locus product Pmel17/gp100/Silv/ME20: controversial in name and in function. Pigment Cell Res 18(5):322–336. doi:10.1111/j.1600-0749.2005.00269.x

15. Watt B, van Niel G, Raposo G, Marks MS (2013) PMEL: a pigment cell-specific model for functional amyloid formation. Pigment Cell Melanoma Res 26(3):300–315. doi:10.1111/pcmr.12067

16. Cao Y, Cao R, Veitonmaki N (2002) Kringle structures and antiangiogenesis. Curr Med Chem Anticancer Agents 2(6):667–681

17. Theos AC, Truschel ST, Tenza D, Hurbain I, Harper DC, Berson JF et al (2006) A lumenal domain-dependent pathway for sorting to intralumenal vesicles of multivesicular endosomes involved in organelle morphogenesis. Dev Cell 10(3):343–354. doi:10.1016/j.devcel.2006.01.012

18. Watt B, van Niel G, Fowler DM, Hurbain I, Luk KC, Stayrook SE et al (2009) N-terminal domains elicit formation of functional Pmel17 amyloid fibrils. J Biol Chem 284(51):35543–35555. doi:10.1074/jbc.M109.047449

19. McGlinchey RP, Shewmaker F, McPhie P, Monterroso B, Thurber K, Wickner RB (2009) The repeat domain of the melanosome fibril protein Pmel17 forms the amyloid core promoting melanin synthesis. Proc Natl Acad Sci USA 106(33):13731–13736. doi:10.1073/pnas.0906509106

20. McGlinchey RP, Yap TL, Lee JC (2011) The yin and yang of amyloid: insights from α-synuclein and repeat domain of Pmel17. Phys Chem Chem Phys 13(45):20066–20075. doi:10.1039/c1cp21376h

21. McGlinchey RP, Gruschus JM, Nagy A, Lee JC (2011) Probing fibril dissolution of the repeat domain of a functional amyloid, Pmel17, on the microscopic and residue level. Biochemistry 50(49):10567–10569. doi:10.1021/bi201578h

22. McGlinchey RP, Jiang Z, Lee JC (2014) Molecular origin of pH-dependent fibril formation of a functional amyloid. ChemBioChem 15(11):1569–1572. doi:10.1002/cbic.201402074

23. Pfefferkorn CM, McGlinchey RP, Lee JC (2010) Effects of pH on aggregation kinetics of the repeat domain of a functional amyloid, Pmel17. Proc Natl Acad Sci USA 107(50):21447–21452. doi:10.1073/pnas.1006424107

24. McGlinchey RP, Shewmaker F, Hu KN, McPhie P, Tycko R, Wickner RB (2010) Repeat domains of melanosome matrix protein Pmel17 orthologs form amyloid fibrils at the acidic melanosomal pH. J Biol Chem 286(10):8385–8393. doi:10.1074/jbc.M110.197152

25. Hu KN, McGlinchey RP, Wickner RB, Tycko R (2011) Segmental polymorphism in a functional amyloid. Biophys J 101(9):2242–2250. doi:10.1016/j.bpj.2011.09.051

26. Kajava AV (2001) Review: proteins with repeated sequence—structural prediction and modeling. J Struct Biol 134(2–3):132–144. doi:10.1006/jsbi.2000.4328

27. Kajava AV, Squire JM, Parry DA (2006) Beta-structures in fibrous proteins. Adv Protein Chem 73:1–15. doi:10.1016/S0065-3233(06)73001-7

28. Yoder MD, Jurnak F (1995) Protein motifs. 3. The parallel beta helix and other coiled folds. FASEB J 9(5):335–342

29. Kajava AV, Steven AC (2006) Beta-rolls, beta-helices, and other beta-solenoid proteins. Adv Protein Chem 73:55–96. doi:10.1016/S0065-3233(06)73003-0

30. Lazo ND, Downing DT (1998) Amyloid fibrils may be assembled from beta-helical protofibrils. Biochemistry 37(7):1731–1735. doi:10.1021/bi971016d

31. Downing DT, Lazo ND (1999) Molecular modelling indicates that the pathological conformations of prion proteins might be beta-helical. Biochem J 343(Pt 2):453–460

32. Lazo ND, Downing DT (1999) Fibril formation by amyloid-beta proteins may involve beta-helical protofibrils. J Pept Res 53(6):633–640

33. Hennetin J, Jullian B, Steven AC, Kajava AV (2006) Standard conformations of beta-arches in beta-solenoid proteins. J Mol Biol 358(4):1094–1105. doi:10.1016/j.jmb.2006.02.039

34. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S et al (2004) UniProt: the universal protein knowledgebase. Nucleic Acids Res 32:D115–D119. doi:10.1093/nar/gkh131

35. Szklarczyk R, Heringa J (2004) Tracking repeats using significance and transitivity. Bioinformatics 20(Suppl 1):i311–i317. doi:10.1093/bioinformatics/bth911

36. Marcotte EM, Pellegrini M, Yeates TO, Eisenberg D (1999) A census of protein repeats. J Mol Biol 293(1):151–160. doi:10.1006/jmbi.1999.3136

37. Biegert A, Soding J (2008) De novo identification of highly diverged protein repeats by probabilistic consistency. Bioinformatics 24(6):807–814. doi:10.1093/bioinformatics/btn039

38. Heger A, Holm L (2000) Rapid automatic detection and alignment of repeats in protein sequences. Proteins 41(2):224–237. doi:10.1002/1097-0134(20001101)41:2<224:AID-PROT70>3.0.CO;2-Z

39. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680

40. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25(9):1189–1191. doi:10.1093/bioinformatics/btp033

41. Marsella L, Sirocco F, Trovato A, Seno F, Tosatto SC (2009) REPETITA: detection and discrimination of the periodicity of protein solenoid repeats by discrete Fourier transform. Bioinformatics 25(12):i289–i295. doi:10.1093/bioinformatics/btp232

42. Di Domenico T, Potenza E, Walsh I, Parra RG, Giollo M, Minervini G et al (2014) RepeatsDB: a database of tandem repeat protein structures. Nucleic Acids Res 42:352–357. doi:10.1093/nar/gkt1175

43. Nummelin H, Merckel MC, Leo JC, Lankinen H, Skurnik M, Goldman A (2004) The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. EMBO J 23(4):701–711. doi:10.1038/sj.emboj.7600100

44. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY et al (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinform. doi:10.1002/0471250953.bi0506s15

45. Tanner DE, Chan KY, Phillips JC, Schulten K (2011) Parallel generalized born implicit solvent calculations with NAMD. J Chem Theory Comput 7(11):3635–3642. doi:10.1021/ct200563j

46. Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B et al (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30(10):1545–1614. doi:10.1002/jcc.21287

47. Huang J, MacKerell AD Jr (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem 34(25):2135–2145. doi:10.1002/jcc.23354

48. Han W, Schulten K (2012) Further optimization of a hybrid united-atom and coarse-grained force field for folding simulations: improved backbone hydration and interactions between charged side chains. J Chem Theory Comput 8(11):4413–4424. doi:10.1021/ct300696c

49. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. J Phys Chem B 111(27):7812–7824. doi:10.1021/jp071097f

50. Qi Y, Cheng X, Han W, Jo S, Schulten K, Im W (2014) CHARMM-GUI PACE CG Builder for solution, micelle, and bilayer coarse-grained simulations. J Chem Inf Model 54(3):1003–1009. doi:10.1021/ci500007n

51. Han W, Schulten K (2013) Characterization of folding mechanisms of Trp-cage and WW-domain by network analysis of simulations with a hybrid-resolution model. J Phys Chem B 117(42):13367–13377. doi:10.1021/jp404331d

52. Han W, Schulten K (2014) Fibril elongation by Abeta(17–42): kinetic network analysis of hybrid-resolution molecular dynamics simulations. J Am Chem Soc 136(35):12450–12460. doi:10.1021/ja507002p

53. Heyer LJ, Kruglyak S, Yooseph S (1999) Exploring expression data: identification and analysis of coexpressed genes. Genome Res 9(11):1106–1115

54. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E et al (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26(16):1781–1802. doi:10.1002/jcc.20289

55. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):33–38

56. Glykos NM (2006) Software news and updates. Carma: a molecular dynamics analysis program. J Comput Chem 27(14):1765–1768. doi:10.1002/jcc.20482

57. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22(12):2577–2637. doi:10.1002/bip.360221211

58. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8(1):52–56

59. Holm L, Park J (2000) DaliLite workbench for protein structure comparison. Bioinformatics 16(6):566–567

60. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res 38:545–549. doi:10.1093/nar/gkq366

61. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos

62. Andrade MA, Perez-Iratxeta C, Ponting CP (2001) Protein repeats: structures, functions, and evolution. J Struct Biol 134(2–3):117–131. doi:10.1006/jsbi.2001.4392

63. Jenkins J, Pickersgill R (2001) The architecture of parallel beta-helices and related folds. Prog Biophys Mol Biol 77(2):111–175

64. Collinson SK, Parker JM, Hodges RS, Kay WW (1999) Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol 290(3):741–756. doi:10.1006/jmbi.1999.2882

65. Baxa U, Cassese T, Kajava AV, Steven AC (2006) Structure, function, and amyloidogenesis of fungal prions: filament polymorphism and prion variants. Adv Protein Chem 73:125–180. doi:10.1016/S0065-3233(06)73005-4

66. Choi JH, May BC, Wille H, Cohen FE (2009) Molecular modeling of the misfolded insulin subunit and amyloid fibril. Biophys J 97(12):3187–3195. doi:10.1016/j.bpj.2009.09.042

67. Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R, Meier BH (2008) Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. Science 319(5869):1523–1526. doi:10.1126/science.1151839

68. Tycko R (2014) Physical and structural basis for polymorphism in amyloid fibrils. Protein Sci 23(11):1528–1539. doi:10.1002/pro.2544

69. Richardson JS, Richardson DC (2002) Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci USA 99(5):2754–2759. doi:10.1073/pnas.052706099

70. Bryan AW Jr, Starner-Kreinbrink JL, Hosur R, Clark PL, Berger B (2011) Structure-based prediction reveals capping motifs that inhibit beta-helix aggregation. Proc Natl Acad Sci USA 108(27):11099–11104. doi:10.1073/pnas.1017504108

71. Harper DC, Theos AC, Herman KE, Tenza D, Raposo G, Marks MS (2008) Premelanosome amyloid-like fibrils are composed of only golgi-processed forms of Pmel17 that have been proteolytically processed in endosomes. J Biol Chem 283(4):2307–2322. doi:10.1074/jbc.M708007200

72. Liou YC, Tocilj A, Davies PL, Jia Z (2000) Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. Nature 406(6793):322–324. doi:10.1038/35018604

73. Shamsir MS, Dalby AR (2007) Beta-sheet containment by flanking prolines: molecular dynamic simulations of the inhibition of beta-sheet elongation by proline residues in human prion protein. Biophys J 92(6):2080–2089. doi:10.1529/biophysj.106.092320

74. Kini RM, Evans HJ (1995) A hypothetical structural role for proline residues in the flanking segments of protein–protein interaction sites. Biochem Biophys Res Commun 212(3):1115–1124. doi:10.1006/bbrc.1995.2084

75. Kay BK, Williamson MP, Sudol M (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J 14(2):231–241

76. Thanka Christlet TH, Veluraja K (2001) Database analysis of O-glycosylation sites in proteins. Biophys J 80(2):952–960