Elastin is a self-assembling protein of the extracellular matrix that provides tissues with elastic extensibility and recoil. The monomeric precursor, tropoelastin, is highly hydrophobic yet remains substantially disordered and flexible in solution, due in large part to a high combined threshold of proline and glycine residues within hydrophobic sequences. In fact, proline-poor elastin-like sequences are known to form amyloid-like fibrils, rich in $\beta$-structure, from solution. On this basis, it is clear that hydrophobic elastin sequences are in general optimized to avoid an amyloid fate. However, a small number of hydrophobic domains near the C-terminus of tropoelastin are substantially depleted of proline residues. Here we investigated the specific contribution of proline number and spacing to the structure and self-assembly propensities of elastin-like polypeptides. Increasing the spacing between proline residues significantly decreased the ability of polypeptides to reversibly self-associate. Real-time imaging of the assembly process revealed the presence of smaller colloidal droplets that displayed enhanced propensity to cluster into dense networks. Structural characterization showed these aggregates were enriched in $\beta$-structure but unable to bind thioflavin-T. These data strongly support a model where proline-poor regions of the elastin monomer provide a unique contribution to assembly, and suggest a role for localized $\beta$-sheet in mediating self-assembly interactions.

Elastin is the extracellular matrix protein responsible for the elasticity of vertebrate arterial vessels, connective tissues, lung and skin. Elastin resilience is afforded by the formation of insoluble fibres, the first step of which involves the self-alignment of elastin monomers, tropoelastin, in a temperature-induced transition known as coacervation, through the association of hydrophobic domains (1,2). Although highly (>75%) non-polar in character, tropoelastin remains predominantly monomeric and structurally disordered in solution (3-5) and critically, retains substantial backbone hydration and flexibility even when assembled (6) and cross-linked into mature, polymeric elastic arrays (7-13).

Maintenance of structural heterogeneity and dynamics of the elastin backbone is achieved in large part by a high proportional composition of both proline (P) (14%) and glycine (G) (35%) residues within hydrophobic sequences (6,14). These residues are commonly arranged into repeated motifs based on recurring sequence elements such as PGV and GVA. This includes a seven-fold tandem PGVGVA repeat in domain 24 of the native human elastin sequence (15). Contrary to enhancing coacervation or elastomeric properties, the multiple substitution of glycine residues for prolines within tandem repeat sequences (e.g. PGVGVA to GGVGVA) results in the formation of amyloid-like fibrils, rich in $\beta$-structure, from solution (6,14). These structures are conformationally more restricted than native elastomeric sequences as a result of the tighter packing of residues into extended secondary structures and concomitant reduction of backbone hydration (6). Interestingly, human elastin domains 28 and 30, which are glycine-rich and relatively proline-poor, form similar amyloid-like fibrils in isolation (16-18).

Examination by Rauscher and colleagues of the proline and glycine content of a wide range of elastomeric and amyloidogenic sequences revealed that a combined PG proportion of (2P+G) > 0.6 correlates with elastomeric assembly (6). From a functional perspective, an increase in conformational order is generally associated with greater fibre stiffness, as evidenced by hydrated spider silks of decreasing proline content (19-21). These data suggest that sequence elements and structural motifs play important roles in modulating
alignment and assembly propensities, and that precise regulation of PG composition is crucial for generating fibres with the required combination of elasticity and strength.

Here we investigated the effect of proline content and spacing on the structure and self-assembly of elastin-like polypeptides by varying the number and arrangement of the fundamental sequence elements PGV and GVA within selected domains. Using a combination of spectroscopic and microscopic techniques we monitored effects on self-assembly, and studied the morphology and architecture of maturing coacervates. These data clearly demonstrate the importance of proline residue distribution within hydrophobic elastin sequences in promoting elastomeric self-assembly, and suggest that relatively proline-poor regions of the native elastin sequence may make a unique contribution to the self-assembly and physical properties of the resulting polymeric elastin.

**Experimental Procedures**

**Polypeptide sequences and domain structures.** The elastin-like polypeptide EP20-24-24 was derived from the native human elastin sequence and contains domain 20 and two copies of domain 24, both hydrophobic domains, separated by cross-linking domains 21 and 23 in the arrangement 20-21-23-24-21-23-24 (22). All polypeptides used in this study were modelled after the domain structure of EP20-24-24. Specifically, variations in the number and arrangement of PGV and GVA motifs were introduced either into the second or both copies of domain 24 to produce hydrophobic domains with altered proline content and/or periodicities (Table 1).

Domains P9, P12 and P6N were named for their proline residue periodicity while domain C20 was modelled on the native tripeptide repeat sequence of domain 20 of chicken elastin. All polypeptides were named after their hydrophobic domain composition and/or interspersed cross-linking sequences (Table 2).

**Kits and reagents.** Restriction enzymes, DNA ligase and T4 polynucleotide kinase were obtained from New England Biolabs and Fermentas. Polymerase chain reactions (PCR) were performed using a HiFidelity Hot Star kit (Qiagen) according to the instructions of the manufacturer. All PCR and ligation products were separated and identified by gel electrophoresis (1.5% agarose) and extracted using a gel extraction kit (Qiagen).

**Plasmid construction.** Oligonucleotides and primers used in plasmid construction were obtained from The Centre for Applied Genomics (The Hospital for Sick Children, Canada). Sequences are provided in Supplemental Material B.

Sequences were constructed from an EP20-24-24 DNA template (22), in which the second or both copies of exon 24 were replaced by sets of oligonucleotides coding for the desired sequence. Further details are provided in Supplemental Material C.

**Polypeptide expression and purification.** Expression and purification of elastin-like polypeptides from *E. coli* BL21 cultures was carried out as previously described (14,22). All peptides were purified by FPLC and their sizes verified by MALDI mass spectrometry (Advanced Protein Technology Centre (APTC), The Hospital for Sick Children). Purity and yields were determined by amino acid analysis (APTC).

**Circular dichroism (CD) spectropolarimetry.** CD spectra were obtained for elastin-like polypeptides (0.2 mg/mL) in water over the wavelength range 250 nm to 190 nm at room temperature. Data were expressed as molar ellipticity [\(\theta]\), calculated as (ellipticity x mean residue weight) / (pathlength (cm) x concentration (mg/mL) x 10).

**Fourier transform infrared (FT-IR) spectroscopy.** Elastin-like polypeptides were dissolved in D\(_2\)O to a concentration of 10 mg/mL and placed between CaF\(_2\) windows for FT-IR analysis. Spectra were obtained using a Nexus 670 FT-IR spectrophotometer (Nicolet Instrument Corporation, USA) over the wavelength range 4000 cm\(^{-1}\) to 650 cm\(^{-1}\). Lyophilized monomeric forms of these elastin-like polypeptides were also analyzed by FT-IR in solid form using an attached microscope (Nicolet continuum) upon deposition on a CaF\(_2\) plate. Similarly, FT-IR data was obtained for aggregates of EP20-P9-P9, EP20-24-P12, EP20-P12-P12, EP20-24-P6N and EP20-P6N-P6N. To induce aggregation, these polypeptides (25 \(\mu\)M) were dissolved in coacervation buffer and heated as described for coacervation until maximum turbidity was reached. Aggregated polypeptide was purified by cycles of gentle pelleting (3000...
rpm, 30 sec) at room temperature and washing with water, and the final suspension was lyophilized. All FT-IR traces were the average of 128 scans with a resolution of 2 cm⁻¹ and were corrected for background. Fourier self-deconvolution was applied to the amide I band of each trace using OMNIC software.

Coacervation. Coacervation was measured as light scattering (turbidity) at 440 nm using a UV-2401PC spectrophotometer (Shimadzu, Mandel, Canada) with attached electronic temperature controller (Shimadzu). Coacervation curves for elastin-like polypeptides (25 µM) were obtained in coacervation buffer (50 mM Tris, pH 7.4) containing 1.5 M NaCl to allow coacervation (turbidity increase) to take place over a measurable temperature range (1,14,22). Temperature was increased at a rate of 1°C/ min, without stirring, to a limiting temperature of 50°C. To monitor the reversibility of coacervation, samples were heated as described until maximal turbidity was reached, then quickly cooled to 20°C at a rate of 5°C/ min. Reversibility was expressed in each case as the decrease in absorbance observed upon cooling to 20°C as a percentage of the maximal absorbance on coacervation.

Light microscopy. Self-assembly of elastin-like polypeptides (100 µM) in coacervation buffer was monitored in real-time using a Zeiss Axiovert 200 inverted epifluorescence microscope with a temperature-controlled Attofluor cell chamber (Molecular Probes, USA). Sample temperature was kept constant at 35°C, approximately 5°C above the coacervation temperature. All samples were imaged for 1 h, beginning at the time of onset of solution turbidity, in differential interference contrast (DIC) mode. Droplet sizes were measured using Velocity 4 software after manually defining the edges of at least 100 droplets per sample, and reported as average diameter ± standard deviation. Statistical analyses were performed using a student’s t-test.

Scanning electron microscopy. Polypeptide aggregates were prepared by heating polypeptides (25 µM) in coacervation buffer to induce maximum turbidity, as described above. Aggregates were desalted by cycles of gentle pelleting (3000 rpm, 30 sec) at room temperature and washing with water. Suspensions were air-dried onto mica grids and sputter-coated with gold for analysis using a FEI XL30 environmental scanning electron microscope (Advanced Bioimaging Centre, The Hospital for Sick Children).

Thioflavin-T binding assay. Elastin-like polypeptides (2 mg/ mL) were prepared in water and incubated overnight at room temperature. Aliquots (10 µL) of polypeptide were incrementally added to a 1 mL solution of thioflavin-T (3 µM) in potassium phosphate buffer (50 mM, pH 6) and left to equilibrate for 5 min. Fluorescence emission at 482 nm was measured using a PTI fluorimeter (814 Photomultiplier Detection System, LPS-220B lamp power supply, Canada) upon excitation at 450 nm. β-amyloid (1-28) (American Peptide Co., CA, USA) was used as a positive control.

RESULTS

Distribution of proline residues in hydrophobic sequences of human elastin. The primary sequence of the human elastin monomer, tropoelastin, generally consists of alternating hydrophobic and cross-linking domains, with each domain corresponding to the translated product of a single exon. Most cross-linking domains in human tropoelastin contain pairs of lysine residues positioned within a polyalanine span of 8-14 residues (KAAK or KAAAK), referred to as a KA-type cross-linking domain. Less common in human tropoelastin are KP-type cross-linking domains containing lysine pairs in KxxK motifs, in which one x position is occupied by a proline. Proline residues are frequent in both hydrophobic domains and KP cross-linking domains. The general alternating domain character of elastin gives rise to a corresponding alternation of structural motifs across the monomer. Proline-poor KA-type cross-linking domains adopt a partial helical structure (1,4,23), whereas hydrophobic domains occupy an ensemble of conformationally disordered states (4), as dictated by the high proline and glycine content within these sequences (6).

However, the contribution of individual hydrophobic domains to assembly is not equal (2,4,14,22,24). Given the high PG composition (2P+G > 0.6) within a diverse range of elastomeric sequences, including hydrophobic elastin domains (6), and the importance of proline residues in modulating structural properties, we mapped the distribution of proline within each continuous hydrophobic interval of human tropoelastin (Fig. 1). Hydrophobic
sequences throughout the first two-thirds of the monomer (up to and including domain 24) preserve average proline residue spacings of 4.5 to 7.8 residues. In contrast, sequences corresponding to domains 26, 28, 30 and 32+33 exhibit greater mean proline residue spacings of between 8.2 and 16.0 residues, commensurate with decreased proline content.

Effect of proline spacing on the structure of elastin-like polypeptides. To probe the effect of increased proline spacing on structure and assembly propensities we constructed a set of elastin-like polypeptides with increasing ratios of GVA to PGV sequence elements within selected domains. Altering the proline periodicity within one or both copies of domain 24 did not substantially affect the secondary structural features of the polypeptides in solution compared to EP20-24-24. For all polypeptides, circular dichroism (CD) spectropolarimetry showed a dominant negative peak around 200 nm representing disordered secondary structure, consistent with previous descriptions of transiently populated local β-turn and polyproline II (PPII) conformations (4), and a shallow trough at 222 nm indicating a minor contribution from α-helix, likely within cross-linking sequences (Figs. 2A, B).

Deconvolution of the amide I band measured by FT-IR spectroscopy revealed a primary peak around 1643 cm⁻¹ for all polypeptides in solution, characteristic of a predominantly disordered conformation (Fig. 3A) (25). The variable presence of a shoulder centred at 1665 cm⁻¹ further indicates unordered (non α-helix, non β-structure) conformations, potentially PPII structure (26), while weaker peaks around 1673 and 1684 cm⁻¹ are attributed to minor contributions from non hydrogen-bonded motifs such as PPII, and from locally hydrogen-bonded β-turns, respectively (27-29).

On lyophilization, most of these polypeptides remained predominantly disordered, as indicated by the peak at 1658 cm⁻¹ (Fig. 3B) (17,29). Exceptions to this were EP20-P12-P12, EP20-24-P6N and EP20-P6N-P6N, which all contained prominent peaks at 1625 cm⁻¹ usually associated with β-sheet (17,25,27), suggesting that these polypeptides were relatively susceptible to the formation of such β-structures.

Coacervation properties of polypeptides. Coacervation is a rapid, temperature-induced phase separation to a colloidal suspension that is characteristic of many elastin-like polypeptides. The transition temperature for this phase separation (coacervation temperature) is determined by both solvent and polypeptide properties (14,22). Coacervation was monitored as a temperature-induced increase in solution turbidity, measured as absorbance at 440 nm without stirring. The coacervation properties of EP20-24-24 under standard conditions (25 µM polypeptide, 50 mM Tris buffer, pH 7.4, 1.5 M NaCl) have been reported previously (14,22). Briefly, coacervation of EP20-24-24 was characterized by a sharp absorbance increase initiated at ~30°C followed by a plateau in maximal turbidity by 40°C (Fig. 4A). Coacervation temperatures remained essentially identical (~30°C) for all polypeptides except EP20-P6N-P6N, which showed an aberrant turbidity profile with absorbance slowly increasing from ~10°C before reaching a plateau by 40°C. Coacervation curves obtained for EP20-C20-C20 and for polypeptides containing P9 domains (EP20-24-P9 and EP20-P9-P9) remained similar to the EP20-24-24 control. However, increasing the average proline spacing to greater than 9 residues reduced the maximum solution turbidity in each case. Replacement of the second copy of domain 24 by either domain P12 (EP20-24-P12) or P6N (EP20-24-P6N) resulted in an approximately 20% decrease in maximum absorbance compared to EP20-24-24, while a 50% decrease in maximum turbidity at plateau was observed for peptides containing two copies of domains P12 (EP20-P12-P12) or P6N (EP20-P6N-P6N) compared to the control (Fig. 4A).

The temperature-induced turbidity increase for coacervation under unstirred conditions is typically reversible upon immediately cooling the sample to below the coacervation temperature after maximum absorbance has been reached. For example, for EP20-24-24, although there is some hysteresis, absorbance returns to pre-coacervation levels by the time the temperature has fallen to 20°C (Fig. 4B). The extent to which the solution turbidity was reversible to baseline (pre-coacervation absorbance) was quantified for each polypeptide by rapidly cooling freshly coacervated samples to 20°C. Patterns in absorbance decreases observed for coacervated EP20-24-24 and EP20-C20-C20 upon cooling were essentially identical (Fig. 4B), and corresponded to a 98% (EP20-24-24) and 91% (EP20-C20-C20) reversal of turbidity to pre-coacervation absorbance,
expressed as a percentage of the maximum absorbance (Fig. 4C). In contrast, polypeptides with increased proline spacing revealed a propensity for irreversible aggregation, measured as the substantial persistence of absorbance even after cooling the solution (Fig. 4C). The turbidity profile of EP20-24-P9 displayed a partial decrease in absorbance, corresponding to a 40% propensity for reversible self-association with respect to maximum turbidity, which was reduced to a 7% turbidity reversal upon the addition of a second P9 domain (EP20-P9-P9) (Figs. 4B, C). Reduction of solution turbidity upon cooling was less than 5% for each polypeptide containing P12 (EP20-24-P12 and EP20-P12-P12) and P6N domains (EP20-24-P6N and EP20-P6N-P6N) (Fig. 4C).

Maturation of coacervate and aggregates. Coacervation of tropoelastin and elastin-like polypeptides is characterized by the formation of colloidal droplets enriched in the polypeptide (30-33). Droplets typically grow (mature) by coalescence before attaining a stable final diameter. The stabilized droplet size is highly dependent on the protein sequence and concentration, and on the solvent environment, including the presence of other extracellular matrix proteins (30,34). In this study the effect of proline spacing on the growth and aggregation characteristics of coacervate droplets was monitored in real-time by light microscopy. Coacervation of EP20-24-24 was characterized by the formation of spherical droplets, which grew by coalescence to a stable average size of 8.8 ± 1.3 µm by one hour after coacervation (Figs. 5A, B). Decreasing proline spacing by replacing both copies of domain 24 with PGV repeats (EP20-C20-C20) did not significantly affect the final size (average diameter 8.5 ± 1.4 µm, p=0.09 compared to EP20-24-24) or morphology of the droplets (Fig. 5C). However, decreasing the proline content and increasing average proline spacing in hydrophobic domains by replacement of one copy of domain 24 with a P9 domain (EP20-24-P9) resulted in a significant reduction in growth of coacervate droplets, with an average droplet diameter of 5.3 ± 0.7 µm (p<0.001 compared to EP20-24-24) after one hour (Figs. 5D, E). A similar replacement of both domains 24 (EP20-P9-P9) resulted in further restriction of droplet growth to 1.9 ± 0.2 µm diameter (p<0.001) compared to EP20-24-P9, and the promotion of droplet clustering (Figs. 5F, G).

Elastin-like polypeptides with further increases in proline spacing were generated by increasing the ratio of GVA to PGV elements, resulting in a proline spacing of 12 residues (P12). Substitution with a single P12 domain (EP20-24-P12) limited the growth of droplets to 1.1 ± 0.1 µm in diameter after 1 hour (p<0.001 compared to EP20-P9-P9) (Figs. 6A, B), with greater clustering into a dense networks of droplets. Scanning electron microscopy of EP20-24-P12 aggregates showed that droplets remained predominantly spherical in shape, even when in contact with neighbouring droplets (Fig. 6C), suggesting the formation of stable and more rigid surfaces inhibiting coalescence. Droplets formed from a polypeptide containing two P12 domains (EP20-P12-P12) were significantly smaller (0.7 ± 0.1 µm, p<0.001) compared to EP20-24-P12 (Figs. 6D, E). Moreover, clustered EP20-P12-P12 droplets displayed more irregular morphologies consistent with partial coalescence events (Fig. 6F).

The sequence of domain P6N was designed to disrupt the spacing of proline residues while keeping the proline content consistent with domain P12. That is, proline residues in domain P6N were confined to the N-terminal half of the domain with a periodicity of 6 residues, leaving a proline-free 29-residue C-terminal sequence of tandem GVA repeats. Such irregular spacing within this P6N domain further retarded droplet growth (~0.5 µm in diameter) compared to polypeptides containing P12 domains (Figs. 7A, B, D, E). Droplet clustering was more pronounced, including apparent assembly into string-like networks, particularly upon the addition of two copies of domain P6N (Figs. 7D, E). Scanning electron microscopy revealed that droplets in these networks were characterized by a partial loss of spheroid morphology in preference for more elongated structures punctuated by residual indentations, indicative of partial coalescence of droplets (Figs. 7C, F). This tendency was further accentuated in networks formed from elastin-like polypeptides containing two P6N domains.

Characterization of aggregates. Given the amyloidogenic propensity of proline-poor elastin-like sequences (6,14,17), the protein aggregates formed by polypeptides with increased spacing between proline residues were tested for the presence of amyloid-like structure. In each case the insoluble aggregate did not contain the cross-beta structure typical of
amyloid fibres, as evidenced by a lack of binding to thioflavin-T dye (data not shown). However, FT-IR analysis of lyophilized aggregates revealed the presence of increased β-content compared to monomeric polypeptide in solution (Fig. 8). Deconvolution of amide I bands revealed a peak at 1625 cm⁻¹ with a less intense peak around 1698 cm⁻¹, both indicative of β-sheet structure (17,25,27). A secondary peak at 1658 cm⁻¹ may be assigned to α-helix and/ or contributions from disordered conformations (17,29), while minor components around 1643 and 1684 cm⁻¹ are representative of disordered conformations and β-turns, respectively (25,27).

DISCUSSION

In this study we present a systematic analysis of the contribution of conserved and repetitive elastin sequence elements PGV and GVA to the structure and self-assembly of elastin-like polypeptides. Decreasing the number of proline residues or increasing their spacing within hydrophobic domains produced a significant deviation from the normal coacervation process, including impairment of the reversibility of coacervation and an increased propensity for the formation of insoluble aggregates. Furthermore, increased proline spacing altered the morphology and growth of colloidal droplets formed on coacervation, progressively resulting in smaller droplets clustering into dense aggregates rich in β-strucutre. These data strongly suggest that closely periodic proline residues within highly hydrophobic elastin sequences are essential for supporting normal self-assembly into elastomeric networks, and preventing the formation of more tightly packed hydrophobic aggregates. This is consistent with the maintenance of structural disorder and flexibility within the full-length monomer that is required for elastic function.

Proline prevents the aggregation of exposed hydrophobic sequences. The folding of globular proteins is driven by the minimization of free energy towards the most thermodynamically stable state, which principally involves the interfacing of polar residues with solvent and the complementary shielding of non-polar residues by burial into a core (35,36). Formation of a core requires the optimal close-packing of hydrophobic residues into regular extended secondary structures, such as α-helix and β-sheet, commensurate with the formation of backbone self-interactions and the exclusion of hydrating water (37). Consequently, adopting a globular tertiary structure is described as a primary mechanism for preventing protein aggregation (38,39). However, it is becoming accepted that hydrophobic sequences may remain exposed concomitant with natively disordered or flexible conformational states, where aggregation is prevented through a combination of evolutionary mechanisms (39). These include selection against the close positioning of residues with high propensity to form tightly packed β-structure (39-44), and the strategic punctuation or (“gate-keeping”) of such sequences by structure-breaking charged or proline residues.

In the case of elastin, the predominantly non-polar sequence contains a high (2P+G = 0.65) proportion of rigid proline residues and flexible glycines (6). Together these residues limit the formation of extended secondary structure by preventing extensive backbone self-interactions, due to the steric constraints (fixed phi dihedral angle and absence of amide hydrogen) imposed by prolines and the high entropic penalty for glycine confinement (6,14). Hence, these residues provide the basis for the maintenance of structural disorder and dynamics within elastin sequences (6,45). Hydrophobic elastin sequences, including tandem PGV and GVA sequence elements, transiently adopt local secondary structures such as PPII motifs and β-turns (4,26,46). The juxtaposition of proline and glycine within recurring PG-based motifs supports the formation of β-turns, with PG forming the corners of the hydrogen-bonded turn (47,48). These turns are labile and impart dynamic flexibility upon “sliding” between successive PG-containing repeats (28).

Analysis of proline residue distribution within human elastin revealed an average punctuation by proline at every 4.5 to 7.8 residues within contiguous hydrophobic sequences between the N-terminus (domain 2) and domain 24. The 6-residue proline periodicity present in domain 24 of EP20-24-24 is consistent with that average, and is clearly sufficient to prevent any tendency for the formation of insoluble aggregates (1,14,22). EP20-C20-C20, containing prolines spaced with a 3-residue periodicity within the C20 sequences, also reversibly coacervated with a
coacervation temperature essentially identical to EP20-24-24, and displayed colloidal droplet behaviour consistent with normal coacervation and elastomeric assembly (49), similar to EP20-24-24. Thus, it does not appear that closer (<6 residue) average spacing of proline residues enhances the formation or maturation of the coacervate. Any additional structural consequences of the extensive tandem PGV repeat found in domain 20 of chicken elastin remain unclear.

Proline distribution within domain P9 was identical to that of native human elastin domain 26, which contains four tandem nonapeptide pseudo-repeats of XGXGXGXPX (X=A, L, I) for a 9-residue proline periodicity. While the coacervation temperatures for EP20-24-P9 and EP20-P9-P9 are essentially identical to that of EP20-24-24, previous studies of a similar elastin-like polypeptide containing domain 26, EP20-24-26, demonstrated that this sequence coacervates at a lower temperature (24°C) than EP20-24-24 (14,22). This may be attributed to the presence of a single arginine residue near the N-terminus of domain 26, which lowers the hydropathy of the sequence. In agreement with its solvent exposed location within the full-length elastin monomer (24), this residue is likely to confer protection from aggregation on the basis of charge and high conformational entropy (50,51).

The strong tendency of polypeptides containing P12 or P6N domains to aggregate is consistent with the previously described contribution of hydrophobicity and solvation effects to peptide aggregation (52). It is known that monomer and aggregated poly(GVA) chains are more hydrophobic and less solvated than chains of poly(PGV) (6). Furthermore, reduction of PG content to below a critical threshold increases the potential for hydrophobic collapse (6). Compared to EP20-P12-P12, EP20-P6N-P6N displayed a highly aberrant coacervation profile, with turbidity appearing at low temperatures and increasing over a wide temperature range, in addition to more restricted droplet growth. These polypeptides share identical sequence composition and mean hydropathy, but differ in the placement of proline residues. In particular, domain P6N contained a proline-free GVA repeat of 29 consecutive residues. While both of these polypeptides exhibited the presence of β-structure as dried monomers as well as in the aggregated state, given the disruptive nature of prolines against β-sheet formation it is probable that domain P6N permits a more extended β-structure motif than domain P12. Moreover, the introduction of greater spacing between proline residues correlated with the progressive retardation of droplet growth, consistent with more rapid loss of surface fluidity and stabilization of droplet surfaces, preventing coalescence (30,34). Growing evidence shows that water-hydrophobic interfaces promote the self-organization of amphipathic sequences into β-sheet structures (53-55). Thus, the presence of longer uninterrupted proline-free sequences in domain P6N could promote increased formation of β-structures at the interface of the colloidal droplets with water, contributing to surface stability and facilitating droplet clustering through direct interactions between β-structures.

Previous studies demonstrate that elastin-like polypeptides containing poly(GVA), poly(GGPGV) or poly(GGPGVA) motifs in place of domain 24 in an EP20-24-24 context form amyloid-like fibrils in solution (6,14). However, the branched, string-like assemblies of small droplets that formed here during the maturation of elastin-like polypeptides with either P12 or P6N domains were clearly comprised of micron-scale globular aggregates, and showed no evidence of amyloid-like supramolecular structure on the basis of their morphology and by lack of binding to thioflavin-T dye. Thus, while both P12 and P6N domains promote the formation of β-structure, either the limited inter-proline spacing in P12 domains or the upstream proline residues in P6N domains preclude the formation of the extensive cross-β structures characteristic of classical amyloid.

Influence of proline spacing on polymeric elastin assembly. The role for the high proportion of proline residues in elastin and other elastomeric proteins is likely related to its function as a “gatekeeper” residue, limiting the extent of sequence able to form β-structures, maintaining the disordered structure required for elastomeric proteins and preventing collapse into hydrophobic cores (6,21,45). Maintenance of an average proline periodicity of approximately 6 residues across hydrophobic domains 2 to 24 of human tropoelastin appears to be consistent with that type of gatekeeper role. However, beyond domain 24 proline spacing in hydrophobic domains of human tropoelastin increases, and the data presented here demonstrate that elastin-like polypeptides allowing such increased
proline spacing show both aberrant coacervation behaviour and the capability of forming at least limited β-sheet structure. Others have also reported altered coacervation behaviour associated with increased β-sheet in elastin-like polypeptides corresponding to sequences downstream of domain 24 (24), and in full-length tropoelastin upon mutation of a single proline to glycine in exon 26, extending a proline-free region (2).

Domain 30 provides a particular example of an extended, repeated hydrophobic sequence with low proline content. In mammalian tropoelastins, domain 30 includes a variable number of GGLG(V/A) repeats with proline-free spans ranging from approximately 16 residues in human to over 50 residues in rat and mouse tropoelastins. The corresponding domain in avian tropoelastin, domain 31, contains a GGLG(V/A) repeat of more than 20 residues. Similar repeated sequence motifs (e.g. GGLGY) are present in a variety of other self-assembling matrix proteins, including dragline spider silks and eggshell membrane proteins of insects, as well as in lamprin, a structural protein of lamprey cartilage. Beta-sheet conformations adopted by these sequences have been suggested to be important for interactions responsible for normal assembly and stabilization of these fibrillar matrix proteins (56,57). Furthermore, while extended regions of β-sheet forming cross-β structures have been associated with diseases of protein misfolding (58), interactions between β-sheets are an important element in the normal folding of globular proteins. Thus, it is possible that interactions involving β-sheet conformations spanning the relatively proline-poor regions near the C-terminus of tropoelastin may also play a role in the assembly of the elastic matrix. Indeed, using a cell culture model of elastic fibre assembly, Kozel et al. (18) demonstrated that deletion of the exon 30 sequence of human elastin prevented the incorporation of this mutated tropoelastin into the elastic matrix, and proposed that the propensity of this sequence for formation of β-sheet structures may provide an essential contribution to elastic fibre assembly.

The data presented here support this proposal. Introduction of longer proline-free regions into hydrophobic domains of elastin-like polypeptides not only affected their coacervation properties, but also correlated with appearance of β-sheet structure and the restriction of growth and increased clustering of colloidal coacervate droplets. Since growth of coacervate droplets takes place through a process of coalescence, cessation of droplet growth is likely related to stabilization of the droplet surface, limiting further coalescence (30,34). Surface stabilization can be the result of interactions with other matrix-associated proteins, as we have shown previously (30), or through the intrinsic ability of coacervates of elastin-like polypeptides to self-organize into fibrillar arrays (34,59). Such organization may be particularly facilitated by propensity for localized β-sheet formation and the preferential partitioning of β-sheet structures at the water/ non-polar interface of the colloidal droplet.

Furthermore, interactions between β-sheet structures arrayed at the surface of droplets could also be a factor contributing to their clustering. In this respect, it is interesting to note that coacervate droplets of full-length tropoelastin, which includes domain 30, are small and show a significant tendency to cluster (31,60). Moreover, in vitro imaging of hydrogel formation after coacervation of human tropoelastin has described network formation through clustering and partial coalescence of coacervate spheres (61). Thus, formation of organized fibrillar structures from coacervates of tropoelastin, in vivo or in vitro, likely requires modulation of surface properties of coacervate droplets, limiting complete coalescence but allowing some fusion and interdigitation of droplet surfaces. Such modulation would be achieved both through interactions with other matrix-associated proteins, as well as by the fundamental colloidal properties of the droplets as determined by protein sequence and structure.

REFERENCES

FOOTNOTES

*The authors thank Ming Miao and Emily Won for construction of EP20-C20-C20, Eva Sitarz for purification of EP20-24-24, Doug Holmyard for SEM images, and Christopher Yip and Jocelyne Verity for FT-IR technical assistance. This project was supported by grants to F.W.K. from the Heart and Stroke Foundation of Ontario (operating grant T6725) and the Canadian Institutes of Health Research (operating grant MOP84496). L.D.M. was the recipient of a Heart and Stroke/ Richard Lewar Centre of Excellence Fellowship.

FIGURE LEGENDS

Fig. 1. Proline residue distribution within hydrophobic sequences of human elastin. Mean proline residue spacing was calculated for continuous hydrophobic regions across the elastin monomer. Designated hydrophobic regions are underlined on the elastin sequence. Domains spanned by each region are marked below the sequence. The human elastin sequence does not contain domains 34 and 35. Lysine residues potentially contributing to cross-link formation are bolded.

Fig. 2. CD spectra of elastin-like polypeptides. Polypeptides were dissolved to 0.2 mg/ mL in water. A. EP20-24-P9 (dashed), EP20-24-P12 (thin line), EP20-24-P6N (dotted). B. EP20-C20-C20 (dotted bold), EP20-P9-P9 (dashed), EP20-P12-P12 (thin line), EP20-P6N-P6N (dotted). In each case the EP20-24-24 control is shown (bold line).

Fig. 3. FT-IR amide I bands of monomeric elastin-like polypeptides. Spectra were collected on samples: A. in solution (10 mg/ mL in D2O) and B. in lyophilized form.

Fig. 4. Effect of proline residue periodicity on the coacervation propensities of elastin-like polypeptides. A. Coacervation of elastin-like polypeptides (25 µM in 50 mM Tris, pH 7.4, 1.5 M NaCl) was measured as an increase in solution turbidity by spectrophotometry (440 nm) upon increasing the temperature (1°C/ min). B. Reversibility of solution turbidity was monitored upon cooling the samples in A to below the coacervation temperature. The decrease in solution turbidity of representative profiles is shown as dashed lines. The profile of EP20-P9-P9 is equivalent to polypeptides containing P12 and P6N domains. C. The extent of absorbance decrease (reversibility) in each case was expressed as a percentage of maximal turbidity.

Fig. 5. Effect of proline residue periodicity on the self-assembly of elastin-like polypeptides. Polypeptides (100 mM) were dissolved in coacervation buffer (50 mM Tris, pH 7.4, 1.5M NaCl) and held at constant temperature above the coacervation temperature. Images were taken 60 min after the initiation of coacervation by bright field microscopy. A, B. EP20-24-24. C. EP20-C20-C20. D, E. EP20-24-P9. F, G. EP20-P9-P9.


Fig. 7. Self-assembly characteristics of EP20-24-P6N and EP20-P6N-P6N. A-C. EP20-24-P6N. D-F. EP20-P6N-P6N. Top and middle panels: bright field microscopy of polypeptides (100 mM) in coacervation buffer (50 mM Tris, pH 7.4, 1.5M NaCl). Images were taken 60 min after the initiation of solution turbidity. Lower panels: scanning electron microscopy of lyophilized polypeptide aggregates.

Fig. 8. FT-IR amide I bands of polypeptide aggregates. Formation of aggregates was induced upon dissolving polypeptides (25 µM) in coacervation buffer (50 mM Tris, pH 7.4, 1.5 M NaCl) and
heating for 1 h at constant temperature above the coacervation temperature. Aggregate assemblies were analysed in lyophilized form.
FIGURE 1

Hydrophobic sequence

<table>
<thead>
<tr>
<th>Hydrophobic Sequence</th>
<th>Average Proline Spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-4</td>
<td>70</td>
</tr>
<tr>
<td>7+8</td>
<td>140</td>
</tr>
<tr>
<td>9+10</td>
<td></td>
</tr>
<tr>
<td>11+12</td>
<td>210</td>
</tr>
<tr>
<td>13-15</td>
<td></td>
</tr>
<tr>
<td>16+17</td>
<td>280</td>
</tr>
<tr>
<td>18+19</td>
<td></td>
</tr>
<tr>
<td>20+21</td>
<td></td>
</tr>
<tr>
<td>24+25</td>
<td>490</td>
</tr>
<tr>
<td>26+27</td>
<td></td>
</tr>
<tr>
<td>28+29</td>
<td>560</td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>32+33+36</td>
<td>698</td>
</tr>
</tbody>
</table>
FIGURE 2

A.

B.

[\theta] (deg cm\textsuperscript{2} dmol\textsuperscript{-1})

wavelength (nm)

-12000

-9000

-6000

-3000

0

190 200 210 220 230 240 250

EP20-24-24

EP20-24-P9

EP20-24-P12

EP20-24-P6N

EP20-24-24

EP20-C20-C20

EP20-P9-P9

EP20-P12-P12

EP20-P6N-P6N
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image A](31\mu m)</td>
<td>![Image C](31\mu m)</td>
</tr>
<tr>
<td>![Image B](10\mu m)</td>
<td>![Image C](31\mu m)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image D](31\mu m)</td>
<td>![Image F](31\mu m)</td>
</tr>
<tr>
<td>![Image E](10\mu m)</td>
<td>![Image G](10\mu m)</td>
</tr>
</tbody>
</table>
FIGURE 6
Table 1. Sequences of elastin-like domains modelled after domain 24. Proline residues are indicated in bold. N=N-terminal half; C=C-terminal half. For complete polypeptide sequences see Supplemental Material A.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Repeat sequence*</th>
<th>(2P+G)†</th>
<th>Number of prolines</th>
<th>Proline periodicity (residues)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C20</td>
<td>(PGV)$_{10}$</td>
<td>1.00</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>24</td>
<td>(PGVGVA)$_{7}$</td>
<td>0.67</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>P9</td>
<td>(PGVGAGIVA)$_{5}$</td>
<td>0.55</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>P12</td>
<td>(PGVGAGVAGVAGVA)$_{4}$</td>
<td>0.50</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>P6N</td>
<td>(PGVGVA)$<em>{4}$ (GVAGVA)$</em>{4}$</td>
<td>0.50</td>
<td>N: 4</td>
<td>N: 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C: 0</td>
<td>C: n/a</td>
</tr>
</tbody>
</table>

* Domains contain an N-terminal GLV motif and a C-terminal GAIPG motif except for C20, which contains an N-terminal GAGVPGVGVPVGIVGIV sequence and a C-terminal VPGVG motif.
† (2P+G) content is expressed as a proportion of domain length (6).
Table 2. Properties and domain composition of elastin-like polypeptides. The sequences of common domains 20 and 21/23 are provided below. For complete polypeptide sequences see Supplemental Material A.

<table>
<thead>
<tr>
<th>Designation</th>
<th>Domain composition</th>
<th>Molecular mass (kDa)</th>
<th>Mean hydropathy (Kyte-Doolittle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP20-C20-C20</td>
<td>20-21-23-C20-21-23-C20</td>
<td>17372</td>
<td>0.75</td>
</tr>
<tr>
<td>EP20-24-24</td>
<td>20-21-23-24-21-23-24</td>
<td>16999</td>
<td>0.92</td>
</tr>
<tr>
<td>EP20-24-P9</td>
<td>20-21-23-24-21-23-P9</td>
<td>17113</td>
<td>0.98</td>
</tr>
<tr>
<td>EP20-P9-P9</td>
<td>20-21-23-P9-21-23-P9</td>
<td>17234</td>
<td>1.03</td>
</tr>
<tr>
<td>EP20-24-P12</td>
<td>20-21-23-24-21-23-P12</td>
<td>17314</td>
<td>1.01</td>
</tr>
<tr>
<td>EP20-P12-P12</td>
<td>20-21-23-P12-21-23-P12</td>
<td>17636</td>
<td>1.09</td>
</tr>
<tr>
<td>EP20-24-P6N</td>
<td>20-21-23-24-21-23-P6N</td>
<td>17314</td>
<td>1.01</td>
</tr>
<tr>
<td>EP20-P6N-P6N</td>
<td>20-21-23-P6N-21-23-P6N</td>
<td>17636</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Domain 20: FPGFGVGGGIPGVAGVPGVGGVPVGVPVGISP
Domains 21/23: EAQAAAAAKAAKYGVTPAAAAAKAAAAAQF
Proline periodicity modulates the self-assembly properties of elastin-like polypeptides
Lisa D. Muiznieks and Fred W. Keeley

J. Biol. Chem. published online October 13, 2010

Access the most updated version of this article at doi: 10.1074/jbc.M110.164467

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2010/11/02/M110.164467.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2010/10/13/jbc.M110.164467.full.html#ref-list-1