"Smart" Behavior of Non-Canonical Elastin-Like Polypeptides

by

Felipe García Quiroz

Department of Biomedical Engineering
Duke University

Date:_____

Approved:

_____
Ashutosh Chilkoti, Supervisor

_____
Gabriel Lopez

_____
Stefan Zauscher

Thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the Department of
Biomedical Engineering in the Graduate School
of Duke University

2011

ABSTRACT

"Smart" Behavior of Non-Canonical Elastin-Like Polypeptides

by

Felipe García Quiroz

Department of Biomedical Engineering
Duke University

Date:_____

Approved:

_____
Ashutosh Chilkoti, Supervisor

_____
Gabriel Lopez

_____
Stefan Zauscher

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Department of
Biomedical Engineering in the Graduate School
of Duke University

2011

Copyright by
Felipe García Quiroz
2011

Abstract

Despite decades of research since the discovery of the environmental sensitivity of tropoelastin, only a handful of elastin-inspired polypeptides departing from the canonical VPGXG motif, where X is any amino acid except proline, have been uncovered. Hence, the field of "smart" protein-polymers has evolved mainly through the introduction of innovative molecular architectures. Instead, we decided to explore sequence diversity as a necessary tool to broaden the biomedical and biotechnological utility of these "smart" protein-polymers. Using a new, highly parallel method for the synthesis of repetitive genes, we conducted a systematic study of the sequence constraints of the canonical VPGXG motif by substituting or inserting Alanine residues along this pentapeptide motif, which yielded new pentapeptide and hexapeptide, non-canonical ELP motifs. These studies led to the discovery of new families of hexapeptide motifs with fully reversible phase transition behavior and suggested an unexpected degree of sequence and conformational promiscuity in the canonical motif that hints at the existence a large space of amino acid sequences with intrinsic "smart" behavior. Moreover, this work shed light into the conformational requirements of the phase transition behavior and suggested the possibility to control the assembly of "smart" protein-polymers in a sequence-controlled manner.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge the many contributions of Miriam Amiram to this work. In particular, Miriam devised the genetically-encoded synthesis strategy that was instrumental to the work presented herein, and worked closely with me on the optimization of this method for the synthesis of the highly repetitive, GC-rich protein-polymers that were the focus of this thesis.

1. Introduction

Elastomeric proteins and in particular elastin, have been the subject of extensive investigations aiming at a molecular understanding of the structure-function relationship among these proteins, with special attention to their remarkable mechanical properties, interesting environmental sensitivity —so-called smart behavior— and self-assembly properties. These studies led to the realization of the important role of protein (structural) disorder in the behavior of these proteins (Muiznieks *et al.* 2010)— protein disorder is now considered to be ubiquitous in biology (Tompa *et al.* 2008)—, and the identification of short recurrent peptides among these proteins capable of forming protein-polymers that recapitulate their structural and environmental properties (Tamburro 2009; Urry *et al.* 2010). Seminal work by Urry and collaborators demonstrated the existence of a general, elastic, environmentally responsive motif Val-Pro-Gly-X-Gly found in elastin (Urry *et al.* 1991), where X is any amino acid except proline, which paved the way for the development of elastin-like polypeptides (ELPs) into an important class of "smart" biopolymers for biotechnology and medicine (Simnick *et al.* 2007; MacKay *et al.* 2009). We refer to this pentapeptide sequence as the canonical ELP motif.

After six decades of research since the discovery of the environmental sensitivity of elastin(Adair *et al.* 1951) (that is, of soluble fragments of elastin and its monomer, tropoelastin), almost 4 decades since the identification of the intrinsic thermo-responsive

behavior of the pentapeptide motif VPGVG (Urry *et al.* 1974), and two decades since the generalization of this repeat unit into the canonical ELP motif VPGXG (Urry *et al.* 1991), only a handful of elastin-inspired polypeptides departing from the canonical sequence have been uncovered, namely minor modifications of the canonical motif such as LPGXG and IPGXG, and the repeat unit VPAVG (Kim *et al.* 2010). Recent efforts have made use of complex bioinformatics tools to search for sequence conservation, amino acid patterns, and recurrent motifs among elastins from different species; these studies have hinted at the potential functional role of the PG dipeptide in elastin (He *et al.* 2007). In a more general approach, Rauscher and collaborators have looked at similarities in Proline and Glycine content between a large panel of elastomeric proteins from different species including elastin, resilin, gluten, and silks among others (Rauscher *et al.* 2006). Noteworthy, these studies have failed to identify first principles for the design of elastomeric, "smart' protein-polymers and novel peptide motifs with intrinsic "smart" behavior.

We reasoned that investigating the sequence constraints of the canonical ELP motif would provide a direct route to the discovery of novel "smart" protein-polymers. We synthesized 9 different variants of ELPs, here referred as non-canonical ELPs, by substituting or inserting Alanine residues along the VPGVG motif, and characterized their "smart", phase transition behavior. These studies revealed an unexpected degree

of sequence and conformational promiscuity in the parent peptide motif and yielded new families of "smart" protein-polymers that display fully reversible thermo-responsive behavior, which will provide a new set of stimulus responsive motifs for biomedical and biotechnological applications.

2. Materials and Methods

2.1 Polypeptide synthesis

2.2.1 Materials

Restriction endonucleases, exonuclease I, calf intestinal alkaline phosphatase (CIP) and T4 DNA ligase (400 U/ul) were purchased from New England Biolabs. PfuUltra™ II Fusion HS DNA polymerase (Pfu) was purchased from Stratagene; Go Taq Green Master Mix™ was purchased from Promega; Circligase™ ssDNA ligase was purchased from Epicentre Biotechnologies; and T4 DNA ligase was purchased from Fermentas and used for the ligation of OERCA products. Plasmid DNA was purified using Qiaprep™ spin miniprep kits and PCR products were purified using the Qiaquick™ PCR purification kit from QIAGEN. Double stranded DNA was visualized by electrophoresis on agarose gels (EMD chemicals) prestained with SYBR® Safe (Invitrogen). BL21(DE3) *E. coli* strains were purchased from EdgeBio. All cultures were grown in Terrific Broth (TB) from BioExpress, and expression was induced with isopropyl β-d-thiogalactopyranoside (IPTG) from Gold Biotechnology. Custom oligonucleotides were synthesized by Integrated DNA Technologies, Inc. Proteins were visualized with precast SDS–PAGE gels (BioRad).

2.1.2 Synthesis of repetitive genes by concatemerization

For each construct, forward and reverse oligonucleotides (Table 1) were annealed at 10 μM in 1X T4 DNA ligase buffer (New England Biolabs) by heating for 3 min at 95

°C followed by slow cooling to room temperature for ~3 hours. The concatemerization of all 3 constructs was explored under various conditions including T4 DNA ligases from Invitrogen and New England Biolabs, different concentrations of the monomer genes (1-2 μM) and 0.1 μM of each annealed adaptor as previously optimized (McPherson *et al.* 1996). The final optimized concatemerization conditions for the generation of the oligomer libraries were largely based on the protocol reported by McPherson *et al* (McPherson *et al.* 1996). This protocol consisted of two 1 hour concatemerization reactions of 1 μM gene in the presence of 400 U of T4 DNA ligase (New England Biolabs) and 0.1 μM of either the 5′ or 3′ adaptor (i.e., to avoid limiting the oligomerization process by blocking the concatenation sites), followed by mixing the two reactions and allowing for further ligation for 2.5 h, all performed at room temperature. These concatemers were then ligated into a custom pET24 vector (McDaniel *et al.* 2010). Transformation and screening of the respective clones was performed as described below for OERCA below, except for the use of Kanamycin for clone selection.

2.1.3 Synthesis of repetitive genes by overlap extension rolling circle amplification (OERCA)

An overview of this synthesis strategy, as developed by Amiram and collaborators (Amiram *et al.* 2011), is depicted in Figure 1.

Preparation of single stranded circular DNA templates

Single stranded 5′ phosphorylated synthetic genes were designed using optimal codons for expression in *E. coli* while minimizing dimerization of the single stranded DNA (ssDNA). The monomer (ssDNA) genes were designed to include 5 repeats of the pentapeptides or hexapeptides resulting from the substitution or insertion of Alanine residues along the pentapeptide VPGVG, as described previously (Amiram *et al.* 2011). The DNA and amino acid sequences for the ELP monomers are reported in Table 2.

250 pmol of ssDNA template gene was circularized in a 50 μL reaction volume containing 5 μL 10X buffer, 2.5 μL Circligase enzyme mix, 2.5 μL $MnCl_2$ and 2.5 μL ATP following the recommendation of the manufacturer. The reaction was incubated at 60 °C for 2 h, PCR purified and then incubated with 20 units of Exonuclease I and 1X Exonuclease buffer at 37 °C for 1 h to remove unreacted ssDNA, followed by a final PCR purification step.

Overlap-elongation rolling circle amplification (OERCA)

The OERCA reaction consisted of 150 ng of circular ssDNA template, 10-40 pmol of sense and antisense DNA primers (Table 1), 2.5 mM of a 70% G/C dNTP mixture and 1μl of Pfu polymerase in a volume of 50 μL. The reaction was incubated at 95 °C for 2 min, followed by 30-40 cycles at 95 °C for 20 s, 55°C for 20 s and 72 °C for 30 s, and finalized at 72 °C for 5 min. The product was purified and visualized on a 1% agarose gel. For most non-canonical ELPs, the OERCA products were then further extended in

the absence of primers at an annealing temperature of 60 °C for 10-15 cycles, with a final extension step at 72 °C for 5 min.

Generation of gene libraries from OERCA products

A modified pET-25b+ vector (Novagen) was used to clone the genes encoding ELP oligomers, which served to incorporate a leader sequence (Ser-Lys-Gly-Pro) and a hexahistidine (C-terminal) tag into the cloned genes (Amiram *et al.* 2011). The vector was prepared by digesting 2 μg of the plasmid DNA with 2 μl of SmaI in 1X NEB buffer 4 for 1.5 h at room temperature, followed by enzymatic dephosphorylation with 1 μl CIP for 15 min to 1 h at 37°C (to prevent self-circularization of the vector), and purification by using a column purification kit. The OERCA products were ligated to the vectors using 5 units of T4 DNA ligase, 2 μl PEG-4000, 1X ligation buffer, ~210 ng of OERCA product, ~250 ng of digested vector, and nuclease-free water in a total volume of 20 μl. The ligation mixture was incubated at room temperature for 3 h, and BL21(DE) cells were then transformed with 7 μl of the ligation mixture following the instructions of the manufacturer. The cells were recovered in SOC media while horizontally shaking at 200 rpm at 37 °C for 40-60 min, and were then plated on TB agarose plates containing 0.1 mg/mL ampicillin.

Screening of gene libraries

Positive clones were screened using a variant of colony PCR (cPCR) referred to as directional cPCR that enabled us to determine the direction of the insert after blunt

ligation (Amiram *et al*. 2011). A ~0.5 μl sample of cells from an existing colony was resuspended in a solution containing 12.5 μl GoTaq green master mix, 10 pmol T7-terminator primer, 10 pmol insert-specific forward primer, and nuclease-free water in a total volume of 25 μl. The PCR reaction conditions were: 95 °C for 2 min for initial denaturation, followed by 30 cycles at 95 °C for 30 s, 52 °C for 30 s, and 72 °C for 1 min. The results of directional cPCR were visualized on a 1% agarose gel stained with SYBR® Safe DNA stain. Positive clones (*i.e.,* those with inserts in the right orientation) were identified by the presence of large smears on the DNA gel. The selected clones were grown overnight in 3 mL TB media supplemented with 100 μg/mL ampicillin, and the plasmids were isolated with a miniprep plasmid purification kit. To verify the size of the insert, a restriction digest was performed with 1 μl EcoRI, 1 μl NdeI, 1X EcoRI buffer and ~200 ng of plasmid DNA. The mixture was incubated for 1 h at 37 °C, and inserts were visualized on a 1% agarose gel stained with SYBR® Safe DNA stain.

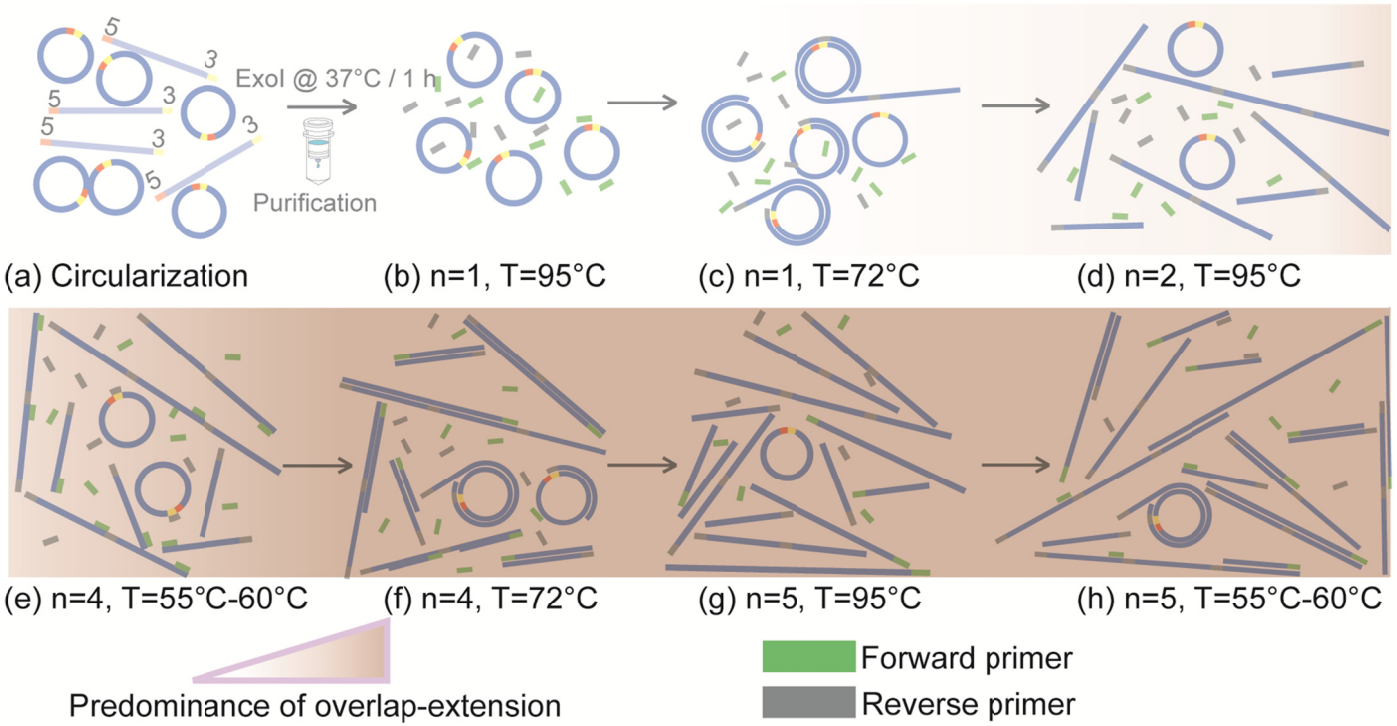

Figure 1: Schematic of snapshots depicting the evolution of an OERCA reaction. (a) A linear oligonucleotide is first circularized yielding a population of circular DNA, which is enriched by removal of the remaining linear DNA before the start of the OERCA reaction (b-h). (b) The circularized oligonucleotide is added to a conventional PCR mixture containing forward and reverse primers. (c) After annealing of the reverse primer, extension in cycle 1 (n=1) occurs primarily in the form of rolling circle amplification. (d) Upon DNA denaturation in cycle 2, linear repeats of the original circularized sequence become available as extension templates. (e-f) As the reactions proceeds, primer annealing and overlap extension preferentially take place to amplify the linear DNA. (g-h) Upon DNA denaturation the repetitive single-stranded DNA is capable of priming/overlapping, which further promotes the extension of the original repeat unit. Reproduced from (Amiram *et al.* 2011).

Table 1: DNA sequence information of the oligonucleotides used to form dsDNA cassettes capable of concatemerization and encoding for five repeats of VPGVA and AVPGVG, and 1 repeat of the glucagon-like peptide-1 analogue (GLP-1).

Repeat unit	Oligomer sequence (5′-3′)	5′ adaptor (5′-3′)	3′ adaptor (5′-3′)
VPGVA	Forward: GTACCAGGTGTTGCTGTGCCGGGCGTGGCCG TGCCTGGCGTGGCGGTCCCGGGCGTCGCGGT TCCAGGTGTAGCA Reverse: TACTGCTACACCTGGAACCGCGACGCCCGGG ACCGCCACGCCAGGCACGGCCACGCCCGGC ACAGCAACACCTGG	Forward: TATGAGCAAAGGT CCC Reverse: TACGGGACCTTTGC TCA	Forward: GTATGGCCGTGATAATCTTCA G Reverse: GATCCTGAAGATTATCACGGC CA
AVPGVG	Forward: GCTGTACCAGGTGTTGGTGCCGTGCCGGGCG TGGGCGCGGTGCCTGGCGTGGGCGCGGTCCC GGGCGTCGGCGCAGTTCCAGGTGTAGGT Reverse: AGCACCTACACCTGGAACTGCGCCGACGCCC GGGACCGCGCCCACGCCAGGCACCGCGCCC	Forward: TATGAGCAAAGGT CCC Reverse: AGCGGGACCTTTGC TCA	Forward: GCTTGGCCGTGATAATCTTCAG Reverse: GATCCTGAAGATTATCACGGC CA

	ACGCCCGGCACGGCACCAACACCTGGTAC		
GLP-1	Forward: CACGTGGAAGGCACCTTTACCAGCGACGTGA GCAGCTATCTGGAAGGTCAGGCGGCGAAAG AATTCATCGCGTGGCTGGTTGTTCCGCGT Reverse: GTGACGCGGAACAACCAGCCACGCGATGAA TTCTTTCGCCGCCTGACCTTCCAGATAGCTGC TCACGTCGCTGGTAAAGGTGCCTTCCAC	Forward: TATGAGCAAAGGT CCC Reverse: GTGGGGACCTTTGC TCA	Forward: CACTGGCCGTGATAATCTTCA G Reverse: GATCCTGAAGATTATCACGGC CA:

Table 2: DNA sequence information for ssDNA templates and primers used for the synthesis of canonical and non-canonical ELPs via OERCA.

Repeat unit	Oligomer sequence (5′-3′)	F primer (5′-3′)	R primer (5′-3′)
VPGVG	GTACCTGGGGTAGGTGTGCCGGGCGTCGGTGTGC CGGGCGTCGGTGTTCCGGGCGTGGGTGTCCCAGG CGTAGGG	GTACCTGGGGTAGGTGT	CCCTACGCCTGGGACA
AVPGVG	GCTGTACCAGGTGTTGGTGCCGTGCCGGGCGTGG GCGCGGTGCCTGGCGTGGGCGCGGTCCCGGGCG	GCTGTACCAGGTGTTG	ACCTACACCTGGAACT

Repeat unit	Oligomer sequence (5′-3′)	F primer (5′-3′)	R primer (5′-3′)
	TCGGCGCAGTTCCAGGTGTAGGT		
VAPGVG	GTAGCTCCAGGTGTTGGTGTGGCCCCGGGCGTGG GCGTGGCGCCTGGCGTGGGCGTCGCGCCGGGCG TCGGCGTTGCACCAGGTGTAGGT	GTAGCTCCAGGTGTTG	ACCTACACCTGGTGCAA
VPAGVG	GTACCAGCTGGTGTTGGTGTGCCGGCCGGCGTGG GCGTGCCTGCGGGCGTGGGCGTCCCGGCGGGCG TCGGCGTTCCAGCAGGTGTAGGT	GTACCAGCTGGTGTTG	ACCTACACCTGCTGGAA
VPGAVG	GTACCAGGTGCTGTTGGTGTGCCGGGCGCCGTGG GCGTGCCTGGCGCGGTGGGCGTCCCGGGCGCGG TCGGCGTTCCAGGTGCAGTAGGT	GTACCAGGTGCTGTTG	ACCTACTGCACCTGGAA
VPGVAG	GTACCAGGTGTTGCTGGTGTGCCGGGCGTGGCCG GCGTGCCTGGCGTGGCGGGCGTCCCGGGCGTCGC GGGCGTTCCAGGTGTAGCAGGT	GTACCAGGTGTTGCT	ACCTGCTACACCTGGAA
APGVG	GCTCCAGGTGTTGGTGCCCCGGGCGTGGGCGCGC CTGGCGTGGGCGCGCCGGGCGTCGGCGCACCAG	GCTCCAGGTGTTGGTG	ACCTACACCTGGTGCG

Repeat unit	Oligomer sequence (5′-3′)	F primer (5′-3′)	R primer (5′-3′)
	GTGTAGGT		
VAGVG	GTCGCTGGTGTCGGTGTTGCCGGCGTGGGCGTGG CGGGCGTGGGCGTAGCGGGCGTTGGCGTTGCAG GTGTAGGT	GTCGCTGGTGTCGGTG	ACCTACACCTGCAACGC CAA
VPAVG	GTTCCAGCTGTTGGTGTGCCGGCCGTGGGCGTGC CTGCGGTGGGCGTCCCGGCGGTCGGCGTACCAG CAGTAGGT	GTTCCAGCTGTTGGTG	ACCTACTGCTGGTACG
VPGVA	GTTCCAGGTGTTGCTGTGCCGGGCGTGGCCGTGC CTGGCGTGGCGGTCCCGGGCGTCGCGGTACCAG GTGTAGCA	GTTCCAGGTGTTGCTGT	TGCTACACCTGGTACCG

2.2 Expression and purification of non-canonical ELPs

Before large-scale expression, 3-5 mL starter cultures of TB media supplemented with 100 μg/mL ampicillin were inoculated with transformed cells from DMSO stocks stored at −80 °C, and incubated overnight at 37 °C while shaking at 200 rpm. The starter cultures were centrifuged at 3000 *g* for 2 min and resuspended in 1 mL of TB medium. Expression cultures (4 L flasks containing 1 L of TB media with 100 μg/mL ampicillin) were inoculated with the resuspended starter culture and incubated at 37 °C with shaking at 200 rpm. After 6-7 h of growth, expression was induced by the addition of IPTG to a final concentration of 1 mM. Cells were harvested 24 h after inoculation, and the polypeptides were purified by inverse transition cycling (ITC) as previously described (Christensen *et al.* 2009).

2.3 Characterization of the phase transition behavior and secondary structure of non-canonical ELPs

To characterize the phase transition behavior of the non-canonical ELPs, the optical density of the polypeptide solutions (50 μM in PBS, unless otherwise indicated in the text) was monitored at a wavelength of 350 nm as a function of temperature, with heating and cooling performed at a rate of 1 °C min^{-1}, on a Cary 300 UV-visible spectrophotometer equipped with a multicell thermoelectric temperature controller (Varian Instruments, Walnut Creek, CA). The derivative of the optical density with

respect to temperature was numerically calculated, and the Tt was defined as the temperature at the maximum of the turbidity gradient.

The secondary structure displayed by non-canonical ELPs was studied by circular dichroism (CD) using an Aviv Model 202 instrument and 1 mm quartz cells (Hellma) by scanning from 280 nm to 180 nm with 1 nm steps and a 3 second averaging time at various temperatures. Purified polypeptides were dialyzed overnight against Milli-Q water, protein purity was assessed by SDS-PAGE and the polypeptides were diluted to 5 μM in water. Raw CD data in millidegrees was first corrected by subtracting the corresponding CD signal from water blanks and transformed into Mean Residue Ellipticity (θ) (Greenfield 2006). Data were considered for analysis whenever the Dynode voltage was below 500 V (Greenfield 2006). The inverse transition temperature of non-canonical ELPs under the conditions specified for CD measurements was also determined as described above.

2.4 Bioinformatics studies

To determine the biological relevance (that is, frequency of occurrence) of non-canonical elastin-like polypeptide motifs in the design of elastins, we studied the amino acid sequence of elastin from multiple species, as listed in Table 3. We retrieved the protein sequence as FASTA (.txt) files from the National Center for Biotechnology Information protein database (http://www.ncbi.nlm.nih.gov/protein), and implemented in-house custom methods in MATLAB software (MathWorks, Natick, MA) to map the

location of canonical (i.e., VPGXG) and non-canonical, hexapeptide elastin-like polypeptide motifs along the sequence of these proteins. To this end, it was assumed that the Ala residue in the non-canonical hexapeptide motifs could be conceived as a non-restricted X residue to yield the non-canonical, general motifs XVPGXG, VXPGXG, VPXGXG and VPGXXG. We did not study XVPGXG since it will inevitably occur with a frequency comparable to that of the canonical motif. The X residue was allowed to take any value expect Pro and Gly. The maps were constructed by digitizing the occurrence of the motifs such that all residues within the motif are assigned a value of 1, and all residues that do not participate in a motif are assigned a value of zero.

Table 3: Sequence information of the different elastin sequences used in the bioinformatics studies.

Protein	Species	Accession Number (GI)
Elastin	Homo sapiens	182021
Elastin	Bos taurus	28461173
Elastin	Mus musculus	31542606
Elastin	Rattus norvegicus	55715827
Elastin	Macaca mulatta	13182892
Elastin b	Danio rerio (Zebrafish)	114326248
Elastin a	Danio rerio (Zebrafish)	121583675

.

3. Results and Discussion

3.1 A highly parallel method for synthesizing DNA repeats enables the rapid synthesis of non-canonical ELPs

The synthesis of multiple structural variants of the canonical ELP motif posed a synthetic challenge, since methods traditionally used for the synthesis of highly repetitive genes are laborious and tend to require recursive cloning steps (Meyer *et al.* 2002). For instance, concatemerization, which is a widely used method to generate tandem repeats of a monomer gene in a single cloning step, has typically failed to generate a large number of repeats of monomer genes encoding for canonical ELPs (McDaniel *et al.* 2010). To assess the feasibility of using concatemerization for the intended studies of the sequence constraints of the VPGVG motif, we chose to study the performance of concatemerization when synthesizing 2 different non-canonical ELPs, a pentapeptide and a hexapeptide, as well as a fairly non-repetitive sequence represented by the GLP-1 gene (Table 1). The concatemerization reactions (i.e., oligonucleotide design, enzyme concentrations, etc.) were selected based on highly successful concatemerization protocols reported in the literature (McPherson *et al.* 1996; McMillan *et al.* 1999; Won *et al.* 2002; Girotti *et al.* 2004). Overall, these different approaches shared the following characteristics in the design of the reactions: (i) 3-nucleotide long overhangs to drive monomer concatemerization (most commonly generated by others by restriction with EarI or a similar restriction enzyme), (ii) concatemerization in the presence of ~400 U T4 DNA ligase (typically from New England Biolabs) for either 12 h

at low temperature (8 °C) or ~3.5 h at room temperature, and (iii) incorporation of capping sequences or adaptors that help reduce concatemer cyclization issues and enable cloning into custom vectors. To design the best possible concatemerization reaction, we opted to directly synthesize dsDNA cassettes with the suggested overhangs so that our reactions did not suffer from issues related to the activity of the selected restriction enzyme. The sequence of the overhangs was allowed to vary among the three genes to avoid inserting extraneous codons into our target polypeptides. The 5′ and 3′ adaptor sequences were almost identical for the 3 constructs, since only the 3 nucleotide overhang needed to vary, enabling the cloning of the concatemerization products into a custom pET24 vector (McDaniel *et al.* 2010).

The concatemerization of these constructs was highly dependent on the selected T4 DNA ligase. For instance, overnight ligation using a T4 DNA ligase from Invitrogen only resulted in simple dimerization or trimerization (data not shown), as has been observed by others (Kurihara *et al.* 2004; McDaniel *et al.* 2010). In contrast, concatemerization in the presence of T4 DNA ligase (400 U) from New England Biolabs under identical conditions resulted in extensive oligomerization (Figure 2a-b). The use of adaptor sequences did not limit oligomerization as shown in Figure 2a, and both overnight and 3.5 h ligation reactions resulted in similar concatemer lengths, which highlights the difficulty in tuning the extent of oligomerization by concatemerization. Despite the degree of oligomerization that was observed, gene libraries constructed from

the concatemerization reactions shown in Figure 2b typically generated libraries with only up to 5 or 6 repeats of the monomer gene (Figure 2c). This performance is comparable to that of overlap-elongation PCR (Amiram *et al.* 2011), but significantly worse than OERCA (Figure 2c). This reduced diversity probably arises from the competitive insertion of smaller concatemers but also from the cyclization of longer concatemers largely favored by the sticky ends required in this type of reaction (despite the addition of adaptors) (Meyer *et al.* 2002). Although the size distribution of the products did not vary for the 3 constructs, the screening of libraries from the most repetitive construct encoding the pentapeptide VPGVA revealed extremely high error rates, which were less pronounced for the hexapeptide AVPGVG and minor for GLP-1 (Figure 2d). The poor performance of concatemerization for the most repetitive sequence revealed that this method is not as useful for GC-rich and highly repetitive DNA sequences. Nevertheless, regardless of the repetitiveness of the monomer gene, concatemerization was shown to be unsuitable, and inferior to OERCA, for the generation of diverse libraries with a large number of repeats. Because of this, we selected OERCA for the rapid synthesis of all the libraries of the canonical and non-canonical ELPs presented herein. A detailed study of the performance of OERCA for the synthesis of non-canonical ELPs was recently published by our group (Amiram *et al.* 2011) and will not be discussed further in this thesis. Overall, we constructed gene libraries encoding non-canonical ELPs that span over a wide range of molecular

weights, as shown in Figure 3, which reinforced the robustness and versatility of OERCA for the genetically-encoded synthesis of highly repetitive protein-polymers (Amiram *et al.* 2011). We also synthesized an additional library encoding polypeptides of the canonical VPGVG motif (data not shown).

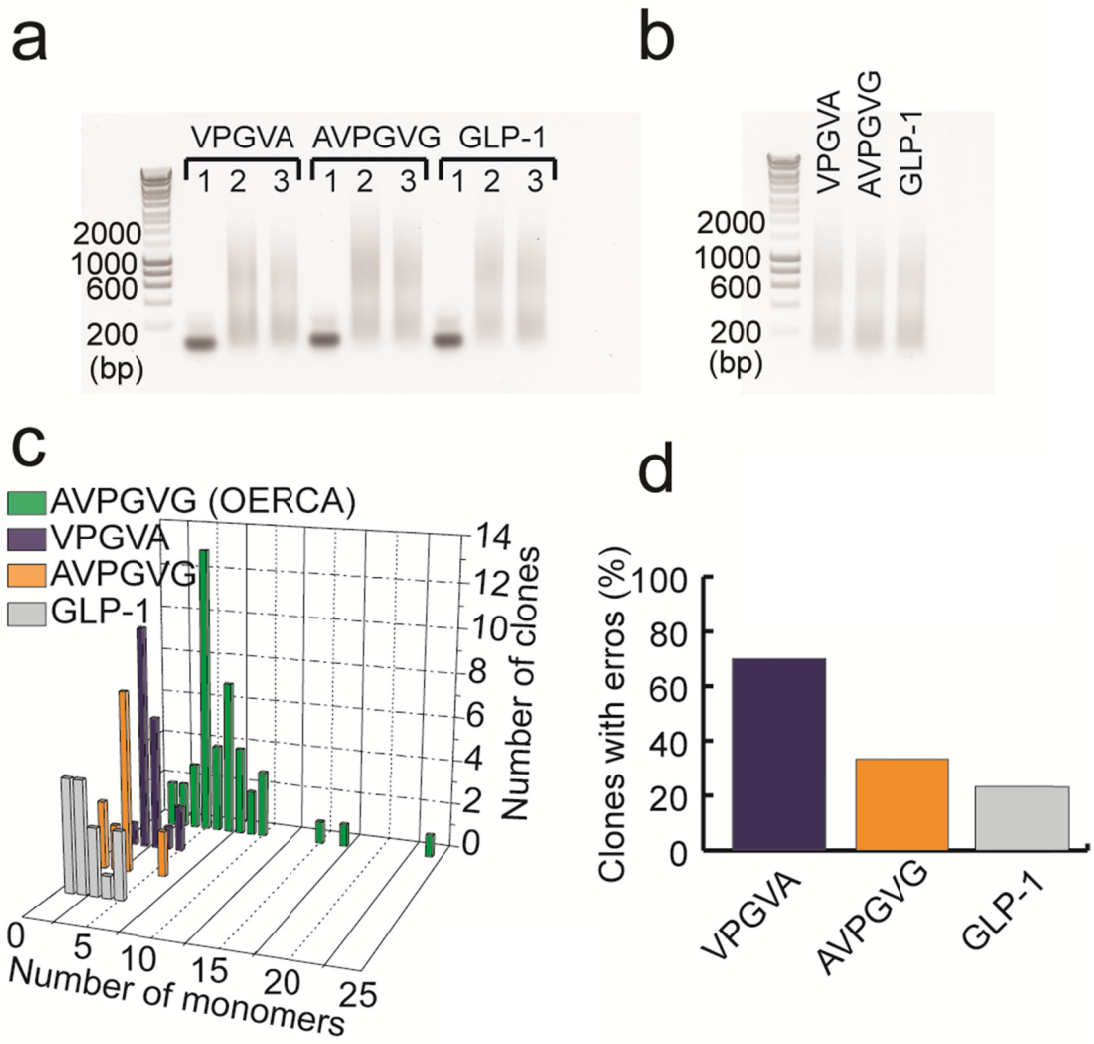

Figure 2: Synthesis of two non-canonical ELP libraries by conventional concatemerization was largely unsuccessful and evidenced the need of using alternative gene synthesis strategies. (a) Annealed monomer genes (1) readily

oligomerized both in the absence (2) and presence (3) of short adaptor cassettes that are required for subsequent cloning steps and that prevent the cyclization of concatemers. (b) Oligomerization for 3.5 h at room temperature in the presence of 400 U of T4 DNA ligase and adaptor cassettes yielded similar oligomerization results. (c) Upon ligation, concatemerization was only able to construct libraries containing short polymers (with up to 5-6 repeats of the monomer gene), whereas OERCA allowed the synthesis of genes with up to 17 repeats. d) GC-rich and highly repetitive sequences (i.e., shorter motifs) proved to be more difficult to prepare by concatemerization, which often suffered from frameshifts and mutations around the concatenation regions. GLP-1 encodes for a non-repetitive short protein that served as control.

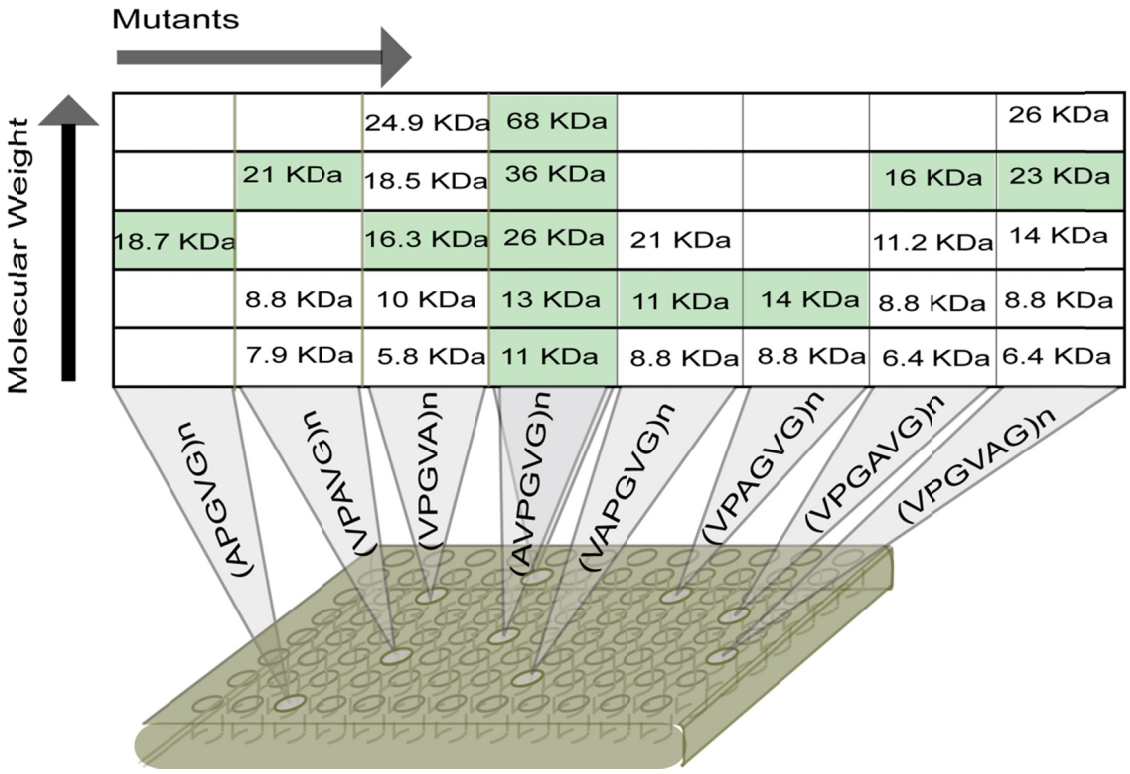

Figure 3: Summary of the gene libraries of non-canonical ELPs constructed by OERCA and indicated by the molecular weight of the encoded protein-polymers. The constructs highlighted in green were used for the subsequent characterization studies. We note that the DNA sequence of all members of these libraries was confirmed by direct DNA sequencing.

3.2 Non-canonical ELPs display "smart", phase transition behavior and shed light into the design of "smart" protein-polymers

The sequence requirements and constraints that govern the stimulus responsive behavior of elastin-like polypeptides remain somewhat of a mystery, despite four decades of investigation. To date, it is unclear to what extent the generalized VPGXG motif (where X is any amino acid except Proline) can be altered without losing thermally responsive coacervation behavior, partly because methods to synthesize genes for ELPs that span a range of compositions and MWs are tedious and difficult to implement in parallel. Motivated by the large gap in our understanding of the phase behavior of ELPs and the increasing interest in these polymers for a wide range of applications (Simnick *et al.* 2007; MacEwan *et al.* 2010), we used OERCA to perform a systematic study of the sequence constraints that govern the phase transition behavior of these protein-polymers.

We generated libraries of genes that encode diverse chain lengths of 9 different Alanine (A) insertion and substitution mutants of poly(VPGVG), as well as the parent motif (Table 1). We refer to these polypeptides as non-canonical ELPs. We screened 96 colonies (of ~200) for each non-canonical motif, which typically yielded ~5 different gene lengths encoding protein-polymers with 10-55 repeats of the pentapeptide or hexapeptide motifs (Figure 3). We then characterized the thermally responsive behavior

of this family of non-canonical ELPs, with the exception of polypeptides with the repeat unit VAGVG, as they expressed poorly.

These experiments resulted in the exciting finding that the entire set of non-canonical ELPs that were characterized display thermally responsive behavior, despite the various mutations introduced to the VPGVG motif (Figure 4a). This finding is significant because it hints at the existence of a large and diverse set of motifs –far larger than the canonical VPGXG motif– that are capable of exhibiting stimulus responsive phase behavior. In previous work by Urry and others (Urry *et al.* 2010) , non-canonical ELP motifs often failed to display fully reversible phase transition behavior with negligible thermal hysteresis, as observed for ELPs (Yamaoka *et al.* 2003; Bochicchio *et al.* 2008). In contrast, we demonstrate, for the first time, the existence of more complex hexapeptide motifs *A*VPGVG, VP*A*GVG, VPG*A*VG and VPGV*A*G capable of displaying fully reversible thermally triggered phase transition behavior (Figure 4b) and environmental sensitivity to both solution temperature and solute concentration.

The thermally-responsive behavior of canonical ELPs is easily and quantitatively tuned by controlling the polypeptide molecular weight (Meyer *et al.* 2004). Therefore, we were interested in studying the molecular weight dependence of the thermally-responsive behavior of these novel non-canonical motifs. We harnessed the ability of OERCA to tune gene length by synthesizing a library encoding for non-canonical ELPs with the hexapeptide motif AVPGVG. This library was assembled from the

transformation reactions of two OERCA products, in which the reaction conditions had been modified to yield two different molecular weight distributions (as depicted in Figure 3). This enabled the construction of a library spanning insert sizes from ~270 to ~2500 bp (Figure 5). We then expressed, purified and characterized the phase transition behavior of 5 constructs spanning the entire range of molecular weights in this library. The transition temperature of these polypeptides decreased in a linear fashion as a function of the reciprocal of the molecular weight, as previously described for canonical ELPs (Figure 5c-d) (Meyer *et al.* 2004). Hence, the thermo-responsive behavior of this non-canonical ELP motif is easily tunable by controlling polypeptide molecular weight. This dependence on molecular weight is also evident in the consistently lower transition temperatures displayed by non-canonical ELPs with increasing number of hexapeptide motifs (Fig. 4b). The functionality –thermal responsiveness– of this library of poly(AVPGVG) spanning polypeptides from ~10 KDa to ~70 KDa also serves to demonstrate that OERCA can be used to readily synthesize repetitive protein-polymers over an unprecedented wide range of sizes without the need for recursive cloning steps.

The behavior of non-canonical ELPs also provided some insight into the role of hydrophobic interactions on the phase transition of elastin and ELPs. Hydrophobic hydration is known to play a major role in determining the phase transition temperature, and the large spread of transition temperatures of non-canonical ELPs with pentapeptide motifs is largely the result of significant changes in the overall

hydrophobicity of the canonical sequence (e.g., replacing a Val by an Ala residue). However, despite having the same amino acid composition (and hence hydrophobicity) and molecular weight, the T_t of polypeptides composed of repeats of AVPGVG and VPAGVG that only differ in amino acid arrangement, are substantially different (Figure 4a). This suggests that the overall hydrophobicity of these polypeptides fails to fully explain their propensity for coacervation (i.e., whether the T_t occurs at high or low temperatures), and in turn suggests that other factors such as their secondary structure are likely to contribute to the differences in their phase transition behavior. This motivated a study of the secondary structure of non-canonical ELPs by circular dichroism (CD), which revealed that the non-canonical ELPS are characterized by an ensemble of highly disordered structures populated by random coil, β-turns and distorted β-sheet conformations (Figure 4c-d), reminiscent of the secondary structure of tropoelastin and ELPs (Yamaoka *et al.* 2003; Bochicchio *et al.* 2008). Interestingly, these studies suggest that different non-canonical ELPs have different secondary structure propensities (Figure 4c-d and Figure 6), as seen differences in their degree of disorder (i.e., the magnitude of the negative peak at ~197 nm) and the fraction of β-turns and/or distorted β-sheets (i.e., negative shoulder around 210 nm) (Yamaoka *et al.* 2003). These differences may eventually lead to a better understanding of coacervation propensity, but we note that this "conformational" promiscuity is observed even among polypeptides that display ideal fully reversible phase transition behavior.

We then investigated the relationship between the phase behavior and secondary structure of two non-canonical ELPs with distinct thermally responsive behavior by concurrently measuring their turbidity profiles and their CD spectra at various stages of the thermally triggered coacervation process (CD spectra were recorded at temperatures indicated by arrows in Figure 4e, g). The non-canonical ELP consisting of repeats of the pentapeptide motif VPAVG exhibited quasi-irreversible phase separation with large hysteresis (Fig. 4e), and complete solubility was only recovered upon cooling the coacervate down to 4 ºC (Figure 7a). In contrast, the non-canonical ELP consisting of repeats of the hexapeptide motif VPGVAG exhibited fully reversible phase transition behavior as a function of solution temperature (Figure 4g). CD spectroscopy illuminated the structural origins of these differences in their thermally triggered coacervation. Whereas both polypeptides showed significant loss of structural disorder as they were heated, as seen by a decrease in intensity of the negative peak at ~197 nm, their CD spectra differed significantly in the mature coacervate stage (i.e., at temperatures corresponding to the maximum absorbance in Figure 4e-, g) and upon subsequent cooling below the T_t. The quasi-irreversible phase behavior of poly(VPAVG) was accompanied by a dramatic loss of structural disorder (i.e., a nearly positive peak at 197 nm) that persisted to a large extent upon cooling well below the T_t (Figure 4f), and the original degree of disorder was only recovered upon cooling to 4 ºC, in accordance with the reversal of the coacervation process as assessed by turbidity data (Figure 7). In

contrast, poly(VPGVAG) exhibited significant residual disorder in the mature coacervate (i.e., negative peak at 197 nm) and showed complete reversal of its secondary structure upon cooling to 25 ºC (Figure 4h) . These results clearly suggest that residual disorder in the mature coacervate and recovery of conformational disorder in non-canonical ELPs are the key to their thermal reversibility. This provides a useful design principle for the engineering of novel "smart" protein-polymers with tunable reversibility.

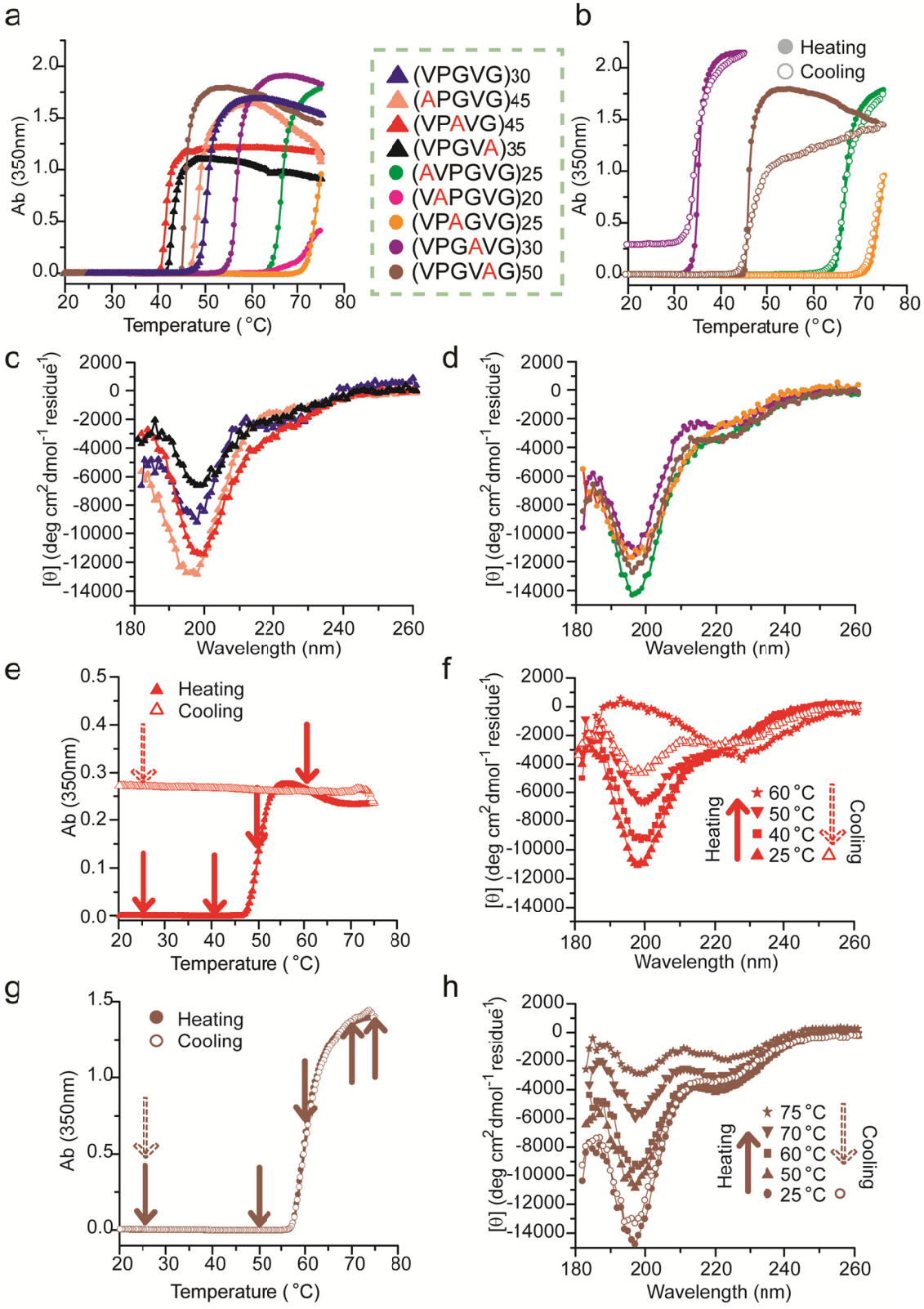
a
b
c
d
e
f
g
h
Ab (350nm)
Temperature (°C)
(VPGVG)30
(APGVG)45
(VPAVG)45
(VPGVA)35
(AVPGVG)25
(VAPGVG)20
(VPAGVG)25
(VPGAVG)30
(VPGVAG)50
Heating
Cooling
[θ] (deg cm2dmol-1 residue-1)
Wavelength (nm)
60 °C
50 °C
40 °C
25 °C
75 °C
70 °C
60 °C
50 °C
25 °C

Figure 4: Thermally responsive behavior of non-canonical ELPs. (a) The turbidity profiles for all non-canonical ELPs exhibit a sharp transition with temperature, characteristic of the inverse phase transition behavior displayed by canonical ELPs. We discovered 4 new hexapeptide motifs (b) that display reversible phase transition behavior. All polypeptides in (a-b) were prepared at a concentration of 50 μM in PBS (i.e., 0.14 M NaCl), except VAPGVG (0.64 M NaCl) and APGVG (2.14 M NaCl) in (a), and VPGAVG (1.14 M NaCl) in (b). Circular dichroism spectra at 25 °C revealed highly disordered conformations predominant in non-canonical ELPs with both pentapeptide (c) and hexapeptide (d) motifs, similar to that of the canonical ELP. (e-h) The coacervation of two non-canonical ELPs with distinct phase transition behaviors was studied by CD at various stages (arrows in e and g) of the coacervation process. The CD spectra and associated turbidity profiles were acquired in water at a polypeptide concentration of 5 μM, and θ indicates the mean residue ellipticity. The pentapeptide motif VPAVG (e-f) underwent quasi-irreversible phase separation (e) and lost its highly disordered conformers (i.e., negative peak at 197 nm) in the mature coacervate and upon cooling to 25 °C (f). In contrast, the hexapeptide motif VPGVAG (g-h) exhibited fully reversible phase transition behavior (g), highly dynamic conformers were preserved in the mature coacervate, and the secondary structure of the polypeptide was completely recovered upon cooling (h). We also confirmed the reversible behavior of VPAVG upon cooling to 4 °C (Figure 7). Reproduced from (Amiram *et al.* 2011).

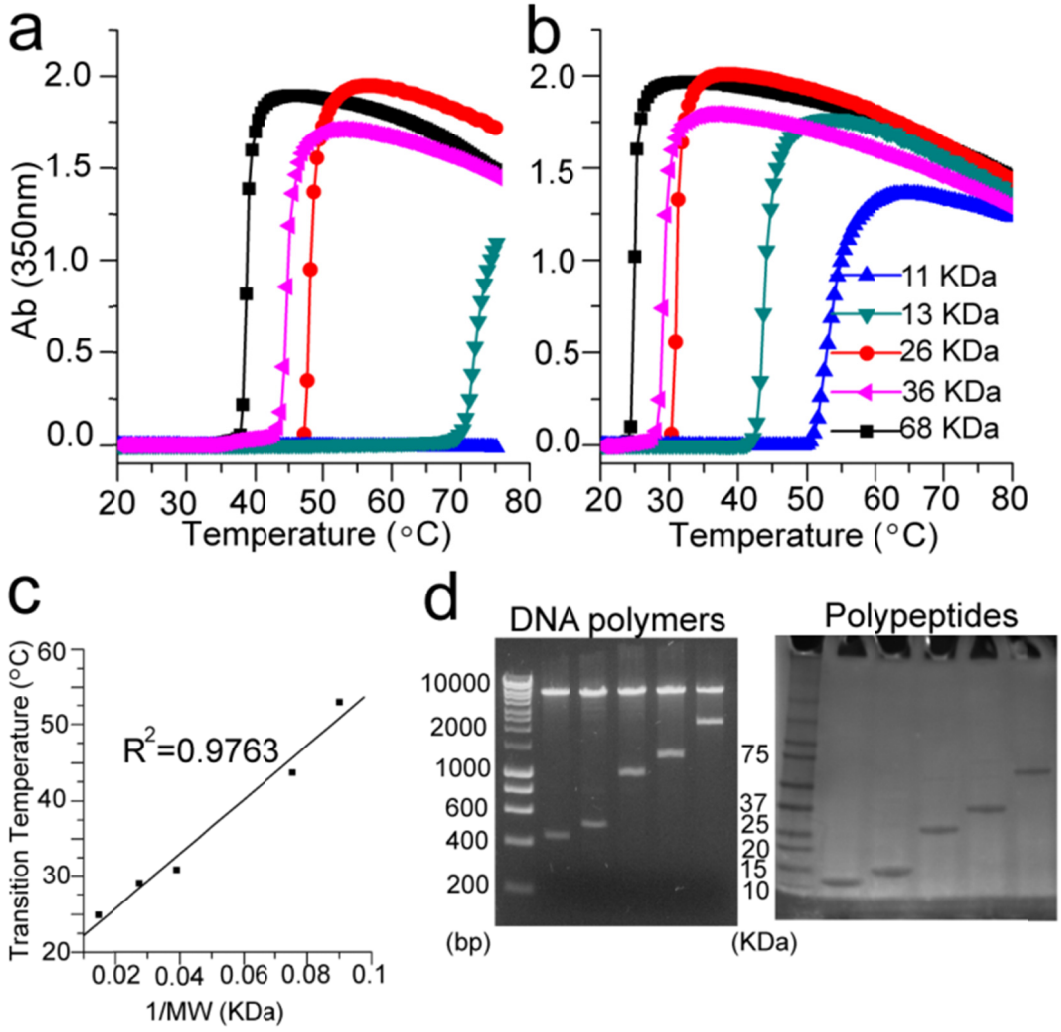

Figure 5: Gene synthesis, expression and characterization of an entire library of the non-canonical ELP with repeating unit AVPGVG. Turbidity profiles for 5 constructs in this library in PBS (a) and PBS supplemented with NaCl to 1.15 M (b). The transition temperatures calculated from (b) varied linearly as the reciprocal of molecular weight of the non-canonical ELP as expected for canonical ELP sequences (c). The wide distribution of molecular weights in this library synthesized by OERCA is illustrated at both the DNA and polypeptide level (d). Reproduced from (Amiram *et al.* 2011).

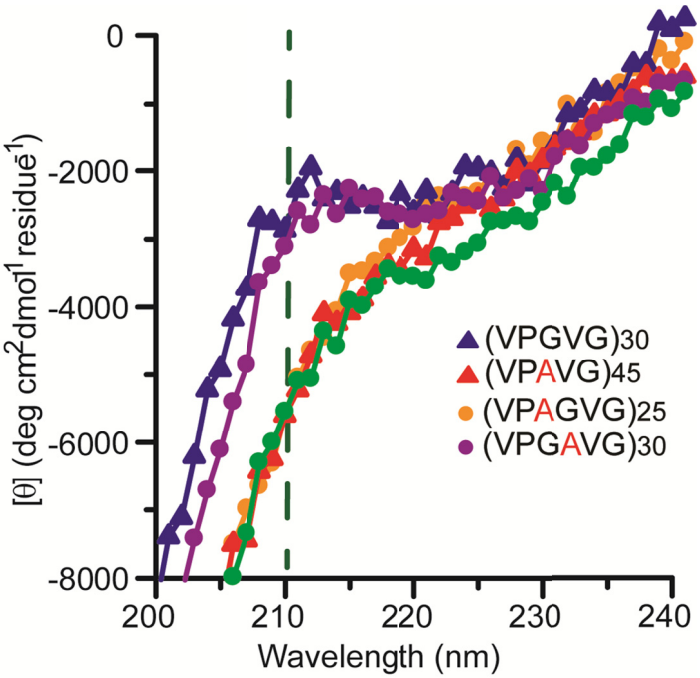

Figure 6: Non-canonical ELPs display an ensemble of highly disordered structures. This figure presents a detailed view of the CD spectra of some non-canonical ELPs below their transition temperature (25 °C). This view underlines the significant differences in the mean residue ellipticity at 210 nm (typically related to β-turn content) and thus points to the promiscuity of allowed conformations within this ensemble of structures.

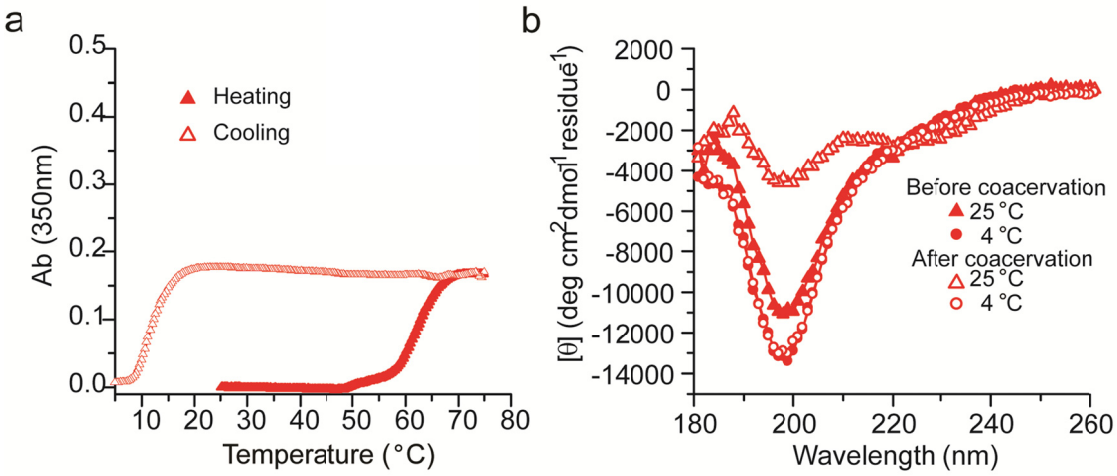

Figure 7: Hysteresis of the quasi-irreversible phase transition behavior displayed by poly(VPAVG). (a) A turbidity profile of poly(VPAVG) as a function of temperature revealed its large thermal hysteresis (> 40 °C), where polypeptide solubility was only recovered upon cooling to 4 °C. (b) CD spectra were acquired at 4 °C and 25 °C before (solid symbols) and after (open symbols) triggering coacervation by heating to 60 °C. The original polypeptide conformation was only recovered upon cooling to 4 °C.

Copolymers of elastin-like polypeptides have been extensively studied to synthesize highly monodisperse, self-assembled nanoparticles (Dreher *et al.* 2008; Simnick *et al.* 2010) and to engineer physical crosslinks into hydrogels and scaffolds (Sallach *et al.* 2009). This assembly is primarily dictated by large differences in the inverse phase transition temperature of individual blocks, such that diblock ELPs wherein one block incorporates hydrophobic guest residues (that is, at position X in VPGXG) and one block incorporates relatively hydrophilic (e.g., alanine or glycine) or even polar (charged or uncharged) residues, would undergo two independent phase separation events at temperatures close to the inverse transition temperatures of the

individual blocks. However, we were interested in asking whether the composition of non-canonical ELPs could be used to control the specificity of their self-assembly. Here we show that the discovery of non-canonical ELPs with new compositions offers an interesting route to refine our control of the self-assembly behavior of ELPs, in a way that is independent of their inverse phase transition temperature. Figure 8 compares the phase transition behavior of mixtures of different canonical and non-canonical ELPs. We studied combinations of protein-polymers that had a nearly constant and small (< 10 °C) difference in their individual phase transition temperatures. Whereas protein-polymers with very similar composition (Figure 8b), with identical or different motif length behaved essentially as a simple combination of two polypeptides of the same motif (Figure 8a), we noted that by changing the composition of the non-canonical ELP motifs we were able to favor a phase behavior with two independent phase separation events (Figure 8c-d). This finding is exciting as it suggests the possibility of driving self-assembly in a sequence controlled manner, which could offer an orthogonal method to drive polypeptide assembly into more complex nanostructures or engineer devices with multiplexing capacity (that is, where one stimulus, say temperature, can have two or more different outputs). However, future studies will be required to address the requirements for achieving the observed specificity in self-assembly. For instance, we observed (data not shown) that non-canonical ELPs with AVPGVG and VPAGVG motifs, which present two distinct amino acid arrangements, lack such specific self-

assembly behavior on mixing and thus undergo a single phase separation event. This may suggest that changes both in motif length —which are related to the periodicity of the polypeptides— and amino acid arrangement are required to achieve such behavior.

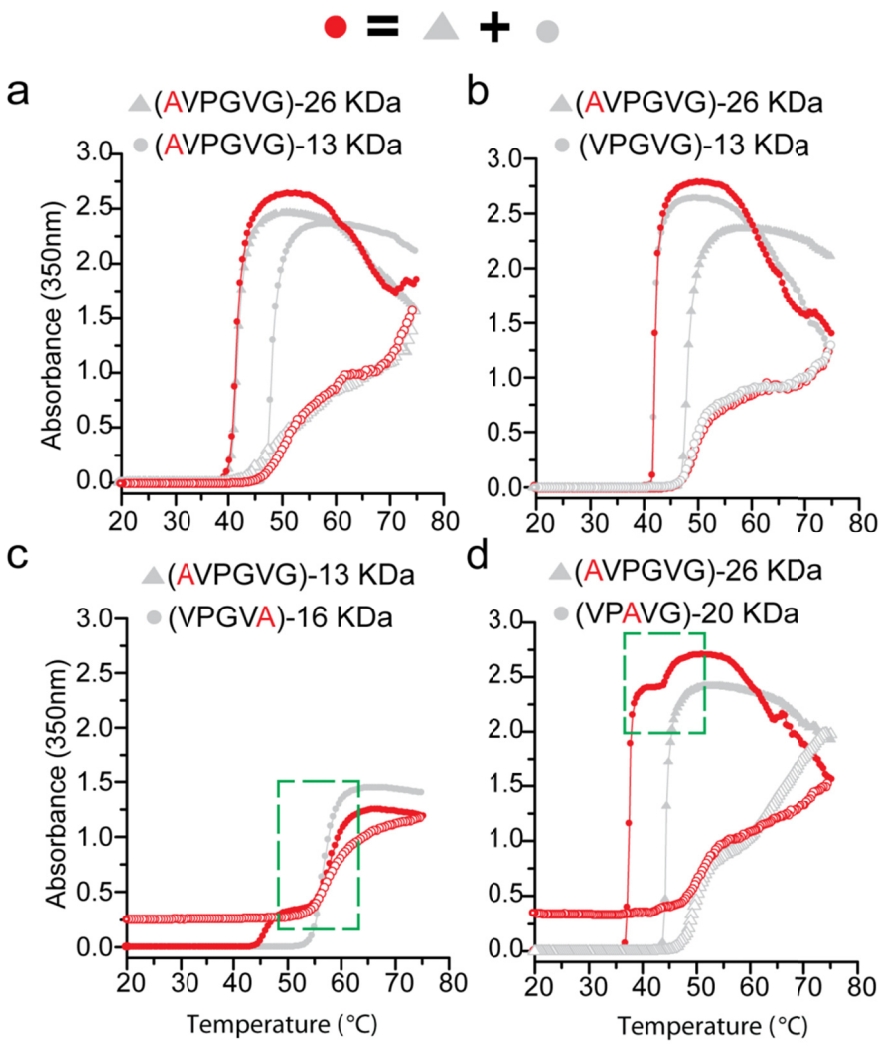

Figure 8: The composition of non-canonical ELPs can be harnessed to promote or prevent co-assembly. Polypeptides with very similar compositions phase separated together (a-b), independently of the difference in motif length (b), whereas non-canonical ELPs with markedly different compositions and length of the repeat unit underwent independent phase separation events (c-d), as demarcated by the green

squares. Polypeptide concentrations were 200 μM (a-b), 50 μM (c) and 100 μM (d) in PBS.

3.3 Non-canonical ELP motifs are rare in natural elastins

The fully reversible phase transition behavior displayed by the non-canonical, hexapeptide ELP motifs (Figure 4), which is analogous to that of the canonical ELP motif, suggested that these motifs might constitute suitable building blocks for the design of elastins. We implemented simple bioinformatics methods to search for generalized forms of the non-canonical motifs along the sequence of elastins from multiple species. Interestingly, we found that these motifs rarely occur in these proteins, unlike the canonical ELP motif (Figure 9). Therefore, it would have been unlikely to arrive at these novel "smart" protein-polymers by conducting traditional bioinformatics studies, which underscores the limitations of the existing paradigm to bio-inspired design that is based on the identification of recurrent and consensus sequences, or copying entire protein domains (Elvin *et al.* 2005; Heim *et al.* 2010). Although the experimental approach undertaken in this study was largely successful in identifying novel compositions of "smart" protein-polymers, we believe that the analysis of the amino acid sequence of elastins and related proteins still offers interesting opportunities for the design of novel "smart" protein-polymers, but that new, unbiased strategies are needed to search for design principles in these proteins, as opposed to specific amino acid compositions.

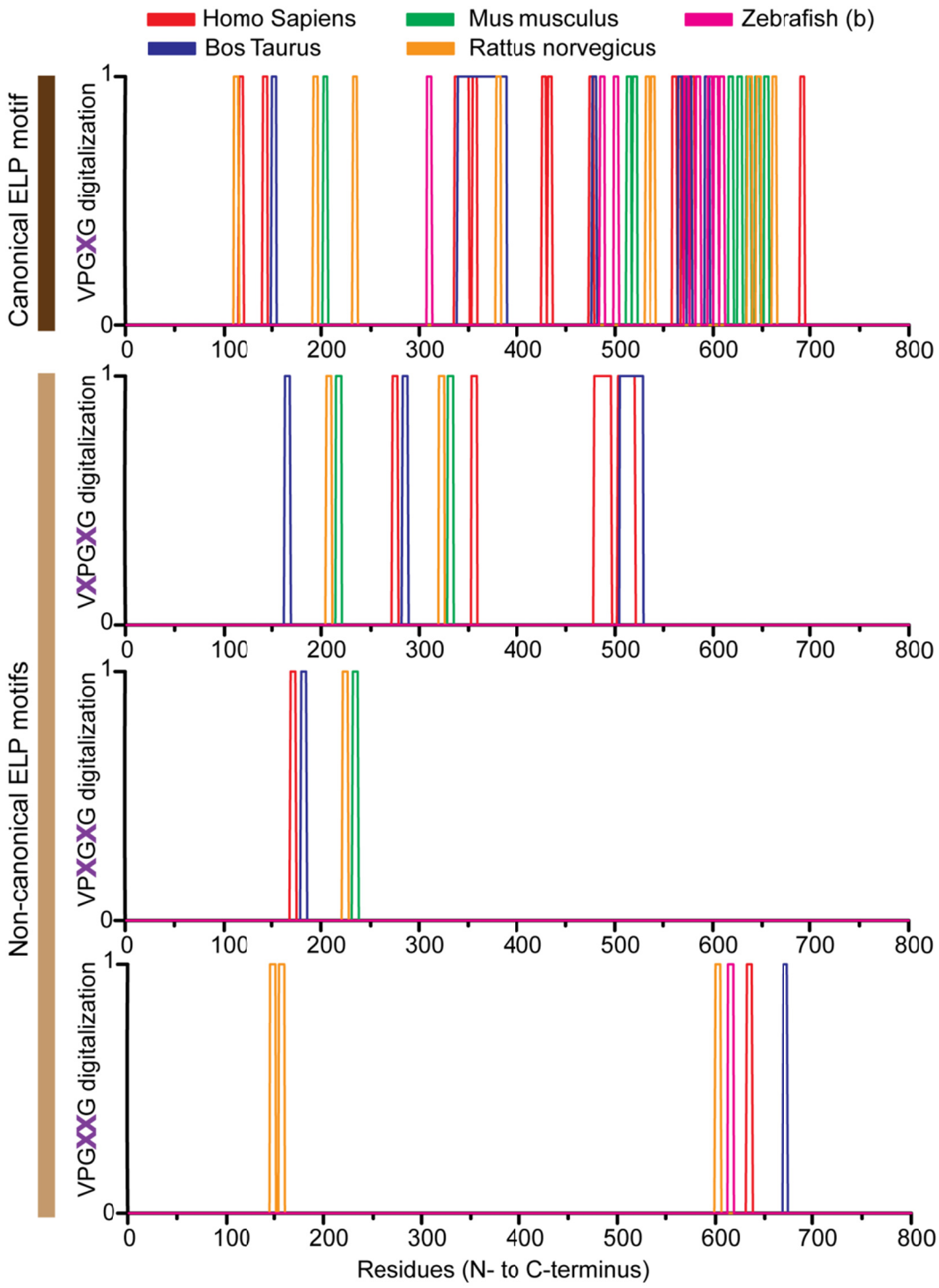

Figure 9: Distribution and abundance of the canonical pentapeptide VPGXG motif and the non-canonical hexapeptide motifs VXPGXG, VPXGXG and VPGXXG

along the sequence of elastin from different species. The localization of the motifs is depicted as a digital signal where all residues forming part of the motif have a value of 1, and 0 otherwise.

8. Conclusion

This work presents three main contributions to the field of "smart" materials: first, we uncovered new families of biopolymers that display stimulus-responsive phase transition behavior analogous to the widely exploited behavior of elastin and elastin-like polypeptides, which may find myriad applications in biotechnology and medicine. Second, our studies of the role of polypeptide conformation in the reversibility of the phase transition behavior evidenced the need for conformational disorder in the design of elastin and other "smart" protein-polymers with fully reversible phase transition behavior; ultimately, this improved understanding may lead to the engineering of "smart" biopolymers with controlled hysteresis in their reversible phase transition behavior by engineering different secondary structure propensities into these biopolymers. Third, we showed the possibility of harnessing the increased diversity in the composition of elastin-like protein-polymers as a means to orchestrate self-assembly in a sequence-dependent manner. Finally, this work hinted at the existence of a large and diverse set of motifs –far larger than the canonical VPGXG motif– that are capable of exhibiting stimulus responsive phase behavior. The exploration of such space of amino acid motifs and the underlying principles that dictate their compositions will be the focus of future studies.

Appendix A

Nature Materials **10**, 141–148 (2011) | doi:10.1038/nmat2942

A highly parallel method for synthesizing DNA repeats enables the discovery of 'smart' protein polymers

Miriam Amiram[a,b], Felipe Garcia Quiroz[a,b], Daniel J. Callahan[a] and Ashutosh Chilkoti[a]

[a]Department of Biomedical Engineering, Duke University, Durham, North Carolina, 27708-0281, USA

[b]These authors contributed equally to this work.

Abstract: Robust high-throughput synthesis methods are needed to expand the repertoire of repetitive protein-polymers for different applications. To address this need, we developed a new method, overlap extension rolling circle amplification (OERCA), for the highly parallel synthesis of genes encoding repetitive protein-polymers. OERCA involves a single PCR-type reaction for the rolling circle amplification of a circular DNA template and simultaneous overlap extension by thermal cycling. We characterized the variables that control OERCA and demonstrated its superiority over existing methods, its robustness, high-throughput and versatility by synthesizing variants of elastin-like polypeptides (ELPs) and protease-responsive polymers of glucagon-like peptide-1 analogues. Despite the GC-rich, highly repetitive sequences of ELPs, we synthesized remarkably large genes without recursive ligation. OERCA also enabled us to discover

'smart' biopolymers that exhibit fully reversible thermally responsive behaviour. This powerful strategy generates libraries of repetitive genes over a wide and tunable range of molecular weights in a 'one-pot' parallel format.

References

Adair, G., H. Davis and S. Partridge (1951). "A soluble protein derived from elastin." Nature **167**: 605.

Amiram, M., F. G. Quiroz, D. J. Callahan and A. Chilkoti (2011). "A highly parallel method for synthesizing DNA repeats enables the discovery of 'smart'protein polymers." Nature Materials **10**(2): 141-148.

Bochicchio, B., A. Pepe and A. M. Tamburro (2008). "Investigating by CD the molecular mechanism of elasticity of elastomeric proteins." Chirality **20**(9): 985-994.

Christensen, T., M. Amiram, S. Dagher, K. Trabbic-Carlson, M. F. Shamji, L. A. Setton and A. Chilkoti (2009). "Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins." Protein Science **18**(7): 1377-1387.

Dreher, M. R., A. J. Simnick, K. Fischer, R. J. Smith, A. Patel, M. Schmidt and A. Chilkoti (2008). "Temperature triggered self-assembly of polypeptides into multivalent spherical micelles." Journal of the American Chemical Society **130**(2): 687-694.

Elvin, C. M., A. G. Carr, M. G. Huson, J. M. Maxwell, R. D. Pearson, T. Vuocolo, N. E. Liyou, D. C. C. Wong, D. J. Merritt and N. E. Dixon (2005). "Synthesis and properties of crosslinked recombinant pro-resilin." Nature **437**(7061): 999-1002.

Girotti, A., J. Reguera, J. C. Rodriguez-Cabello, F. J. Arias, M. Alonso and A. M. Testera (2004). "Design and bioproduction of a recombinant multi(bio)functional elastin-like protein polymer containing cell adhesion sequences for tissue engineering purposes." Journal of Materials Science-Materials in Medicine **15**(4): 479-484.

Greenfield, N. J. (2006). "Using circular dichroism spectra to estimate protein secondary structure." Nature Protocols **1**(6): 2876-2890.

He, D., M. Chung, E. Chan, T. Alleyne, K. C. H. Ha, M. Miao, R. J. Stahl, F. W. Keeley and J. Parkinson (2007). "Comparative genomics of elastin: Sequence analysis of a highly repetitive protein." Matrix Biology **26**(7): 524-540.

Heim, M., C. B. Ackerschott and T. Scheibel (2010). "Characterization of recombinantly produced spider flagelliform silk domains." Journal of structural biology **170**(2): 420-425.

Kim, W. and E. L. Chaikof (2010). "Recombinant elastin-mimetic biomaterials: Emerging applications in medicine." Advanced drug delivery reviews **62**(15): 1468-1478.

Kurihara, H. and T. Nagamune (2004). "DNA polymerase-catalyzed elongation of repetitive hexanucleotide sequences: Application to creation of repetitive DNA libraries." Biotechnology Progress **20**(6): 1855-1860.

MacEwan, S. R. and A. Chilkoti (2010). "Elastin-Like Polypeptides: Biomedical Applications of Tunable Biopolymers." Biopolymers **94**(1): 60-77.

MacKay, J. A., M. Chen, J. R. McDaniel, W. Liu, A. J. Simnick and A. Chilkoti (2009). "Self-assembling chimeric polypeptide–doxorubicin conjugate nanoparticles that abolish tumours after a single injection." Nature Materials **8**(12): 993-999.

McDaniel, J. R., J. A. MacKay, F. G. Quiroz and A. Chilkoti (2010). "Recursive Directional Ligation by Plasmid Reconstruction Allows Rapid and Seamless Cloning of Oligomeric Genes." Biomacromolecules **11**(4): 944-952.

McMillan, R. A., T. A. T. Lee and V. P. Conticello (1999). "Rapid assembly of synthetic genes encoding protein polymers." Macromolecules **32**(11): 3643-3648.

McPherson, D. T., J. Xu and D. W. Urry (1996). "Product purification by reversible phase transition following Escherichia coli expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP." Protein Expression and Purification **7**(1): 51-57.

Meyer, D. E. and A. Chilkoti (2002). "Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: Examples from the elastin-like polypeptide system." Biomacromolecules **3**(2): 357-367.

Meyer, D. E. and A. Chilkoti (2004). "Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides." Biomacromolecules **5**(3): 846-851.

Muiznieks, L. D., A. S. Weiss and F. W. Keeley (2010). "Structural disorder and dynamics of elastin." Biochemistry and Cell Biology **88**(2): 239-250.

Rauscher, S., S. Baud, M. Miao, F. W. Keeley and R. Pomes (2006). "Proline and glycine control protein self-organization into elastomeric or amyloid fibrils." Structure **14**(11): 1667-1676.

Sallach, R. E., W. Cui, J. Wen, A. Martinez, V. P. Conticello and E. L. Chaikof (2009). "Elastin-mimetic protein polymers capable of physical and chemical crosslinking." Biomaterials **30**(3): 409-422.

Simnick, A. J., D. W. Lim, D. Chow and A. Chilkoti (2007). "Biomedical and biotechnological applications of elastin-like polypeptides." Polymer Reviews **47**(1): 121-154.

Simnick, A. J., C. A. Valencia, R. Liu and A. Chilkoti (2010). "Morphing low-affinity ligands into high-avidity nanoparticles by thermally triggered self-assembly of a genetically encoded polymer." ACS nano **4**(4): 2217-2227.

Tamburro, A. M. (2009). "A never-ending love story with elastin: a scientific autobiography." Nanomedicine **4**(4): 469-487.

Tompa, P. and M. Fuxreiter (2008). "Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions." Trends in biochemical sciences **33**(1): 2-8.

Urry, D., M. Long, B. Cox, T. Ohnishi, L. Mitchell and M. Jacobs (1974). "The synthetic polypentapeptide of elastin coacervates and forms filamentous aggregates." Biochimica et Biophysica Acta (BBA)-Protein Structure **371**(2): 597-602.

Urry, D. W., C. H. Luan, T. M. Parker, D. C. Gowda, K. U. Prasad, M. C. Reid and A. Safavy (1991). "Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity." Journal of the American Chemical Society **113**(11): 4346-4348.

Urry, D. W., K. D. Urry, W. Szaflarski and M. Nowicki (2010). "Elastic-contractile model proteins: Physical chemistry, protein function and drug design and delivery." Adv Drug Deliv Rev.

Won, J. I. and A. E. Barron (2002). "A new cloning method for the preparation of long repetitive polypeptides without a sequence requirement." Macromolecules **35**(22): 8281-8287.

Yamaoka, T., T. Tamura, Y. Seto, T. Tada, S. Kunugi and D. A. Tirrell (2003). "Mechanism for the phase transition of a genetically engineered elastin model peptide (VPGIG)(40) in aqueous solution." Biomacromolecules **4**(6): 1680-1685.