Collagen-like block copolymers with tunable design

Production in yeast and functional characterisation

Helena Teles

Thesis committee

Thesis supervisor
Prof. dr. G. Eggink
Professor of Industrial Biotechnology
Wageningen University

Thesis co-supervisor
Dr. F.A. de Wolf
Senior scientist at Food and Biobased Research
Wageningen University

Other members

Prof. dr. W.E. Hennink	Utrecht University
Prof. dr. ir. V. Martins dos Santos	Wageningen University
Prof. dr. ir. W. Norde	Wageningen University
Dr. M. Hendriks	DSM, Geleen

This research has been conducted under the auspices of the graduate school VLAG

Helena Teles

Collagen-like block copolymers with tunable design

Production in yeast and functional characterisation

Thesis

Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M. J. Kropff
in the presence of the
thesis committee appointed by the Academic Board
to be defended in public
on Tuesday 7th September 2010
at 13:30 h in the Aula

Helena Teles
Collagen-like block copolymers with tunable design - Production in yeast and functional characterisation

PhD thesis, Wageningen University, The Netherlands (2010) - 152 pages

ISBN: 978-90-8585-708-2

Free will is doing gladly and freely that which one must do.

C. G. Jung

CONTENTS

Chapter 1 General Introduction 9

Chapter 2 Precision gels from collagen-inspired triblock copolymers 29

Chapter 3 Influence of chain length on gel-forming properties of telechelic collagen-inspired polymers 57

Chapter 4 Hydrogels of collagen-inspired telechelic triblock copolymers for the sustained release of proteins 77

Chapter 5 Pilot-scale fermentation and purification of collagen-inspired triblock copolymers in *Pichia pastoris* 95

Chapter 6 General Discussion 115

Summary 135

Acknowledgements 141

About the Author 147

Chapter 1

General Introduction

1.1 PROTEIN-BASED BIOMATERIALS

When we think about proteins, we easily recognise the important role and variety of functions they play in living organisms. Through millions of years of evolution, nature has created and refined proteins for a wide variety of specific purposes. The sequence of amino acids, the proteins building blocks, determines the three-dimensional folded structure of each protein which is responsible for most of the important protein properties. By using the same set of 20 amino acid building blocks, nature can vary, almost infinitely, the protein structure and thus the functional properties of proteins. The best known role of proteins is as enzymes, which catalyse chemical reactions, but proteins also have a role in the signalling and transducing of signals, and as building materials (structural proteins) in the body.

Many proteins derived from natural sources have interesting properties, and provide materials for a number of applications across many markets[1], for example: silk, a natural protein produced by the mulberry silk worm can be woven into textiles, whey (milk plasma) is used as an emulsifier in the dairy industry, collagen a protein derived from animal bones and tissue and casein, present in cow's milk, have been used as adhesives. However, natural proteins are not susceptible to structural modifications of their architecture for specific needs[2]. Applications are often limited to the existing protein functional properties and available quantities[1].

Artificial protein polymers are of increasing interest in biotechnology[3]. Based on natural protein structures, amino acid sequences can be chosen to create specific folding patterns, and thus, desired material properties. Nature is an endless source of inspiration where material scientists can find basic protein motifs for the design of new kinds of high performance materials. For instance, mussels have the ability to adhere to their support using an adhesive protein which is quite effective in salt water[3]. In the wall of arteries**,** are found elastin and collagen with life spans of the order of one century[3].

However, the capacity to design and synthesise proteins requires a deep understanding of how molecular architectures assemble in nature. Although understanding sequence-to-structure relationship in proteins is challenging, there is a growing number of peptide folding motifs for which the so-called ‘design rules’ are becoming available and understood. These motifs include α-helical coils and the β-structured amyloid-like assemblies and collagens[4].

Peptides and proteins may be produced by chemical or biological processes. Chemical processes allow the effective production of relatively small quantities of custom-made peptides in a short period of time. But the production of sequences over 35 amino acids are generally not considered economically viable[5]. In addition, long peptides are obtained by polycondensation methods, where the chance of introducing sequence errors increases with the length of the peptide. Recombinant methods allow the production of pure (monodisperse) and long proteins (10-200 kDa), and are for this purpose superior to chemical synthesis routes. The development of such systems is, nevertheless, relatively long as compared to chemical synthesis[6] and less effective for the production and purification of short peptides (10-30 kDa)[5]. The advances made in recombinant DNA technology in the past decade have tackled some of the problems inherent to the production of recombinant proteins in biological systems, and contributed to their development. In particular, the use of genetically modified microorganism for the production of recombinant proteins is one of the most promising systems because it offers high productivity and usually scalable and cost-effective processes[5, 7, 8].

1.2 COLLAGEN AND GELATIN AS BIOMATERIALS

Collagens are the most abundant proteins in animal tissue. They constitute approximately one third of mammalian body proteins, and their main function is to provide mechanical strength to different tissues and organs as extracellular, structural proteins. The basic structural motif of collagen is adapted to meet the specialised needs of different tissues such as bone, skin, tendon, teeth, blood vessels, and cartilage[9-11]. So far 29 genetically distinct types of collagen have been identified[12]. They have been divided into several families based on the polymeric structure they form or related structural features[10].

Besides their structural and supportive role, collagens also have a number of other important biological functions. Collagens interact with cells in connective tissue (muscle, tendon, etc.) and transduce essential signals for the regulation of cell anchorage, migration, proliferation, differentiation, and survival[11, 13].

Based on its structural role and compatibility within the body, isolated collagen is a commonly used biomaterial for several medical applications (surgical and pharmaceutical) as well as for some food purposes[6, 14]. The use of collagen for medical applications can be traced back to prehistoric times where collagen in the form of tendons was used as suture material[15]. Gelatin, obtained from denaturated and partially degraded collagen, is also used in a number of different applications in the food and pharmaceutical industries, as well as for some technical applications, mostly in the manufacture of photographic materials. Gelatin has been employed in other technical uses such as in adhesives, electrolytic metal refining, micro-encapsulation and sizing of paper. However, in some of these applications, such as adhesives, gelatin has been replaced by synthetic products[6].

The use of collagen and gelatin as expensive biomaterials has enormous commercial potential[10, 11, 14, 16]. Injectable collagen or gelatin formulations, hemostats for the control of bleeding during surgery or after trauma, sponges for burns/wounds, gel formulations in combination with relevant bioactive molecules for sustained drug delivery, skin and bone substitutes, are only some of the biomedical applications and products found in literature[7, 14, 16-20].

Bovine and porcine tissues are today’s main source of collagen and gelatin. The most abundant type of collagen in animal tissues is collagen type I, and therefore most of the medical preparations contain this type of collagen. But type I collagen is not the only collagen present in animal tissues, a minor, but variable, amount is type III and other collagens[6, 7, 16]. Furthermore, the ratio between the different types of collagen and the degree of crosslinking vary with the type and age of the tissue used to extract collagen. This results in a lack of reproducibility, and heterogeneity between batches[16]. Gelatin is a heterogeneous material as well, composed of a mixture of collagen chains of different length, structure, and composition, which yields products with variable gelling and physical-chemical properties.

The variability in composition and structure of animal-derived collagen and gelatin presents a significant challenge for those using these proteins in medical applications[7] where reliable and predictable materials are essential. However, this is not the only concern related with the use of these materials in medicine. The potential presence of infectious agents, such as virus or prions, in collagen and gelatin derived from animal sources, and the possibility of inducing immunogenic reactions, poses a risk for patients receiving the medical product[7, 16]. Despite these, animal derived collagen and gelatin have been used for decades in medical and pharmaceutical applications.

In recent years, the biomedical relevance of collagen and gelatin and the advances in recombinant-based production systems have motivated scientists to find alternatives to animal collagen and gelatin. The recombinant (microbial) production of collagen and gelatin opens up the possibility of producing these proteins free of animal contaminants and with defined molecular composition and structure, and thus predictable properties. It might also be advantageous to have collagen and gelatin with human amino acid sequences in view of immunogenetic responses. This is only possible using recombinant production since human collagen (from placentae, human cell culture, etc.) is scarcely available and human gelatin, i.e. derived from human collagen, is not available at all[6]. Recombinant production allows the possibility to produce collagen and gelatin with human identical amino acid sequences or to produce fully designed collagen-like proteins, with the aim of creating new and varied structures with novel functionalities which may provide benefits for specific applications. The advances in the recombinant production of collagen and gelatin will be addressed in section 4 of this chapter.

In order to find successful strategies for the biotechnological production of this class of proteins a good understanding of collagen structure and biochemistry is necessary.

1.3 STRUCTURE AND BIOSYNTHESIS OF COLLAGEN

1.3.1 Amino acid sequence and secondary structure

Chemically, collagen is characterised by an exceptionally high content of glycine, proline and hydroxyproline, organised in repeating (Gly-Xaa-Yaa)$_n$ sequences that fold into a triple-helical structure (Figure 1.1) The length (n), the content of proline and hydroxyproline, and the presence of other amino acids in either Xaa or Yaa positions vary with the type of collagen. Collagen molecules also contain non-collagenous domains, and many of these have important biological functions that are different from those of the collagen domains. One example is the non-collagenous domain of type IV collagen which can inhibit angiogenesis and tumour growth[21].

The typical collagen triple helix consists of three extended left-handed proline II-like helical chains that are super coiled into a right-handed triple helix. The inner core of the collagen helix is rather crowded, therefore the presence of glycine (the smallest amino acid, with a side chain consisting only of hydrogen) is required in every third position in order for the polypeptide chain to pack together close enough to form the collagen triple helix[10]. The three chains are staggered by one residue with respect to each other, in the direction of the helix axis, and are linked through interchain hydrogen bonds[6]. The hydrogen donors are the peptide -NH groups of glycine residues, and the hydrogen acceptors are the peptide -CO groups of residues on the X and Y positions[11].

Proline is often found in the X position and hydroxyproline in the Y position. Because of their ring structure, these amino acids stabilise the helical conformations of the polypeptide chains. The unusual amino acid hydroxyproline is formed within the endoplasmic reticulum by modification of proline residues that have been already incorporated into collagen peptide chains. Some lysine residues in collagen are also converted to hydroxylysine. The hydroxyl groups of these modified amino acids and bridging water molecules participate in hydrogen bonds between polypeptide chains, which further stabilises the triple helix[9, 22, 23]. In short, the collagen helical form is the consequence of several reinforcing bonds, each of which is relatively weak[9].

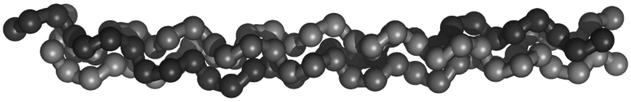

Figure 1.1 Collagen triple-helix

1.3.2 Stability of the collagen helix- the role of proline and hydroxyproline

The thermal stability of the collagen helix is an important biological property. The melting temperature (T_m) of collagens (usually defined as the point where half of the helical structure of a collagen solution is lost) is related to the body temperature of the source species[9, 24], and therefore collagens from different species have different melting temperature. This difference in thermal stability is strongly related to the imino acid content (proline and hydroxyproline) present in the collagen. The higher the imino acid content, the more stable the helix[6, 9, 25]. A correlation between denaturation temperatures and the sum of proline and hydroxyproline content was demonstrated for several vertebrates by Burge and Hynes[26] and by Piez and Gross[27]. Studies of the melting temperature of chemically synthesized polypeptide models of collagen have been used to clarify the biological significance of the high content of proline and hydroxyproline. Peptides made of tandem repeats of Gly-Xaa-Yaa possess an intrinsic propensity to form the collagen-like triple helix. Peptides with the structure (Gly-Pro-Pro)n (n= 10, 15, or 20) were the first ones to be synthesised[28].

Those have shown to form triple-helical structures similar to the triple helical structure of collagen[29]. Later, the peptides (Pro-Hyp-Gly)n (n=5 or 10) were synthesised and were also found to form triple helical structures[30]. Indeed, the presence of hydroxyproline (Hyp) greatly increased the thermal stability of the helices. The midpoint of the thermal transition of helix to coil (T_m) for the (Pro-Hyp-Gly)$_{10}$ was ~60 °C, while that for the peptide (Pro-Pro-Gly)$_{10}$ was ~25 °C, about 35 °C lower. These results confirm the greater stabilising effect of hydroxyproline as compared to proline. Recent studies, have also shown that the melting temperature of partially hydroxylated recombinant collagen is reduced[31], and that non-hydroxylated custom-designed or human gelatins do not form helices at least above 4 °C[32-34]. The T_m also depends upon the length of the collagen peptide chain, as it is expected for a cooperative transition[35]. For example, the melting temperature determined for (Pro-Pro-Gly)$_{15}$ is ~52 °C[29], more than two fold higher than the T_m of (Pro-Hyp-Gly)$_{10}$. Additional factors may be involved in the stabilisation of the triple helix as suggested by other studies performed with collagen-like polypeptides[36-38], such as the unexpected stabilising effect of arginine in the Y position[39].

1.3.3 Collagen biosynthesis

The synthesis of fibrilar collagens in the animal cell is a complex process involving a number of intracellular and extracellular post-translational modifications. Collagens are first synthesised in the rough endoplasmic reticulum as soluble precursor molecules known as procollagen. Procollagen molecules contain non-helical propeptide extensions at both N- and C- terminal ends that are important in the process of chain association and nucleation, and prevent the premature formation of fibres inside the cell[10, 11]. Before chain association and triple helix formation, the procollagen chains are modified inside the lumen of the endoplasmic reticulum, where hydroxylation of certain proline and lysine residues and glycosylation takes place[6, 11]. The assembly and thermal stability of the procollagen molecules is strongly dependent on the successful hydroxylation of proline residues in the Y position into

4-hydroxyproline by the enzyme prolyl 4-hydroxylase[6, 40]. In the absence of this enzyme the procollagen molecules remain in the form of a non-triple helical, non-functional protein[6]. In addition to this enzyme, a number of other proteins that assist in the correct folding and processing of collagen (chaperones and enzymes respectively) are required[40]. The modified procollagen molecules self-assemble from -N to -C terminus into a triple helix in a zipper like manner, and are secreted to the extracellular space, where they fill the space between the cells, and bind cells and tissues together. After secretion, the C- and N- terminus extensions are cleaved off by specific proteinases and the collagen molecules spontaneously self-assemble into fibrils, which are stabilised by the formation of covalent cross-links between the side chains of lysine and hydroxylysine residues. See Figure 1.2 for a schematic representation of collagen biosynthesis in the animal cell.

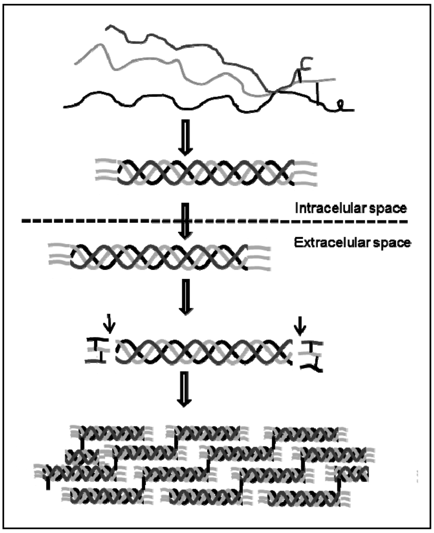

Figure 1.2 Schematic representation of the biosynthesis of collagen in the animal cell

The extent and type of cross-linking varies with the physiological function and age of the tissue[9, 16]. The most abundant type of collagen (type I collagen) is one of the fibril forming collagens that are the basic structural component of connective tissues. Several other types of collagen do not form fibrils, but play other roles in different types of extracellular matrices. For example, fibril-associated collagens, which are present in connective tissue as well, bind to the surface of collagen fibrils and connect them both to one another and to other matrix components. Network-forming collagens, e.g. type IV collagen, are more flexible than fibril-forming collagens because the Gly-X-Y repeats are frequently interrupted by non-helical sequences. As a result, they assemble into two-dimensional crosslinked networks instead of fibrils[10, 22].

1.4 CHARACTERISTICS OF NATURAL GELATIN

Gelatin is typically prepared from collagen type I or III, which is present in skin and bones of cows or pigs. During the conversion of collagen into gelatin, the covalent cross-links formed between the collagen-fibrils, and the hydrogen and hydrophobic bonds formed within the collagen triple-helical structure are destroyed.

The industrial production of gelatin consists of three main stages: pre-treatment of the raw material, extraction of the gelatin, and purification and drying. The pre-treatment consists essentially of either an alkaline (liming) or an acid treatment. The main purpose of liming is to break covalent cross-links in collagen and thus improve the efficiency of extraction of gelatin. Liming is mainly applied to animal tissues in which the collagen has a high degree of covalent bonding. In contrast, no covalent bonds are broken during acid treatment. This type of treatment is usually limited to tissues of young animals or other tissues with a small amount of covalent bonding[6]. Following pre-treatment the gelatin is extracted using hot water. The heat breaks the hydrogen and hydrophobic bonds within the triple-helix structure and leads to its dissociation into individual soluble chains, and small fragments. Upon cooling, the chains rewind in triple-helical structures[41]. However, because the chains have lost their initial register, they do not align as complete collagen triple-helices. Instead, they form interconnected triple helical junction zones in a random fashion. Just like in a fishing-net, the triple helical junctions knot the chains together in a network. The helix formation is determined by the coincidental proximity of polypeptide domains and the potential thermostability of helices formed by interaction of those domains[6]. As a consequence, helices of different amino acid sequence and composition, different lengths, and different intrinsic thermostability are formed[36, 37, 42]. Figure 1.3 shows a schematic representation of the conversion of collagen into gelatin.

Regardless of the process used, gelatin preparations consist of a distribution of polypeptide fragments of different sizes, different isolectric points and different gelling properties[14, 43]. Furthermore, the physical-chemical properties of the final product depend on the extraction method used, the length of the thermal denaturation, and the electrolyte content of the resulting material[6, 14, 17, 43]. The extreme complexity of natural gelatin preparations makes the study of their structure very difficult.

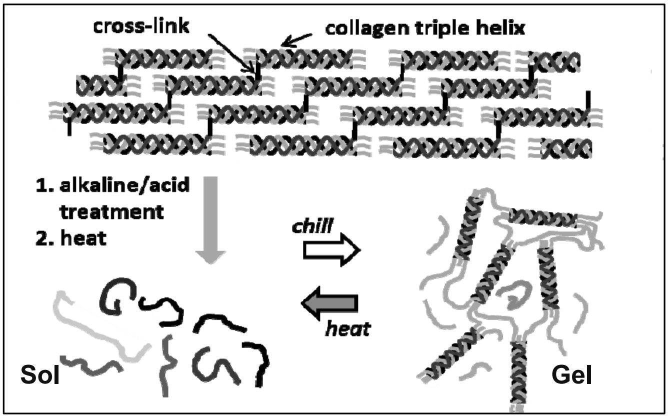

Figure 1.3 Conversion of collagen into gelatin

1.5 RECOMBINANT PRODUCTION- FROM COLLAGEN TO COLLAGEN-LIKE PROTEINS

Efforts have been increasing in the last few years to develop an efficient and safe alternative production system of recombinant human collagen and gelatin. Synthesis of recombinant collagen presents various challenges, not least because of the many post-translational modifications required for a recombinantly produced collagen molecule to achieve a fully folded, triple helical conformation. As mentioned above in section 1.3.2 one of the most important post-translational modifications is the hydroxylation of specific proline residues by the enzyme prolyl 4-hydroxylase.

Several microbial and non-microbial systems have been investigated for the production of recombinant collagen and collagen-like polymers. Mammalian[44, 45] or insect cell lines[46, 47], which should possess appropriate post-translational modification systems, have been used for the production of recombinant human collagen. Mammalian cells transfected with human collagen genes expressed and secreted hydroxylated full-length collagen; while insect cells accumulated the collagen intracellularly and were incapable of producing sufficient collagen-specific proalyl 4-hydroxylase (P4H) to fully hydroxylate the collagen. Recombinant non-hydroxylated or partially hydroxylated collagen has been shown to accumulate in milk from transgenic mice[48] and in secretions from transgenic silkworms[49]. Transgenic tobacco cell culture[14] and transgenic tobacco plants[50, 51] have also been used for the production of recombinant collagen. Though, the collagen expression levels achieved in all these expression systems were too low to make these systems commercially attractive.

High-yield expression systems, such as bacteria and yeast do not contain any P4H activity and can therefore not be used for production of fully hydroxylated collagen without co-expression of a suitable P4H.

Production of hydroxylated and non-hydroxylated collagen-like proteins has been attempted in *Escherichia coli*[52, 53]. But, expression levels were generally low, mainly due to product degradation (in the case of non-hydroxylated molecules), and gene instability issues.

Since yeasts are eukaryotes, they offer certain advantages over bacteria, such as occurrence of most of the post-translational machinery needed for functionality and/or stability of recombinant animal proteins (glycosylation, acylation, methylation, protein folding, targeting to subcellular compartments and proteolytic adjustments); efficient secretion of proteins, which facilitates the isolation of the desired heterologous protein; and ability to cope with highly repetitive coding sequences, which is important in recombinant gelatin and collagen expression[54, 55]. Furthermore, yeasts grow to high cell densities resulting in the production of larger quantities of recombinant proteins than in mammalian and insect cells systems.

Different yeast strains have shown to be able to produce human collagen. However hydroxylated triple-helical collagen secretion is a problem in the yeast species used to date, *Saccharomyces cerevisiae, Pichia pastoris* and *Hansenula polymorpha*[31, 56, 57]. In contrast, single chain collagen-like molecules with natural and synthetic sequences are very efficiently secreted by both *H. polymorpha*[58] and *P. pastoris*[32-34].

Completely folded triple helical collagen is resistant to proteases[59, 60], while most unfolded collagen-like molecules are very sensitive to proteolytic degradation[14, 33, 61]. Among all yeast used so far, the methylotrophic yeasts, *P. pastoris*, appears to be the most suitable for high-yield secreted production of non-hydroxylated collagen and collagen-like proteins, i.e. protein polymers made of repeating (Gly-Xaa-Yaa) sequences. In *P. pastoris*, proteolytic degradation of unfolded collagen-like molecules (gelatin) is much less of a problem than, for instance, in *H. polymorpha*[58]. In addition, the expression level of a non-hydroxylated procollagen fragment in *H. polymorpha* was considerably lower than expression levels of non-hydroxylated collagen-like proteins in *P. pastoris*[58].

1.6 *P. PASTORIS* AS A HOST SYSTEM FOR THE PRODUCTION OF COLLAGEN AND COLLAGEN-LIKE PROTEINS

The efficacy of *P. pastoris* for the production of homogenous, well defined, recombinant collagen and collagen-like proteins has been reported in several works[32-34, 62].

Collagen type III and I and the human P4H genes have been expressed in *P. pastoris* with some success. Although the use of this system requires the integration of several genes in the *P. pastoris* genome and is dependent on the correct association between the co-expressed α and β human P4H subunits[62, 63], moderate levels (0.5 g/l) of triple helical human procollagen type I were accumulated inside the yeast cells[62].

Werten and co-workers[33] reported the use of *P. pastoris* for the production of recombinant non-hydroxylated gelatins based on rat type III and mouse type I collagen sequences. One of the fragments, derived from rat type III collagen, was produced at one of the highest levels for a secreted heterologous protein ever obtained from recombinant yeast (14.8 g/l). High-level expression and secretion of a highly hydrophilic designed gelatin was also successfully achieved in *P. pastoris*. These studies demonstrate the ability of this yeast to express sequences from different origins (natural or artificial). Furthermore, the use of high productivity *P. pastoris*-based recombinant systems results in cost-effective processes for the production of large amounts of material[7], because the medium components are inexpensive and defined, consisting of pure carbon sources (glycerol and methanol), biotin, salts, trace elements, and water. This medium is free of undefined ingredients that can be source of pyrogens or toxins and is therefore compatible with the production of human pharmaceuticals[55].

Some *P. pastoris*-derived recombinant collagen-like products have been evaluated for use in medical devices and pharmaceutical products and shown to have good biocompatibility. Recombinant collagen (type II) sponges, commonly used as scaffold in tissue engineering applications, revealed to be more biocompatible than animal-derived collagen, as determined in animal studies[16]. A low molecular weight recombinant gelatin was developed for use as stabiliser for various injectable biologics, and proven to be safe for use in humans[14, 32]. Olsen et al.[14] reported that recombinant gelatin fragments were able to support attachment of Vero cells[64] (cell line derived from kidney epithelial cells) at levels similar to those provided by commercial bovine type I collagen.

Despite all the research and development around *P. pastoris*-based recombinant collagen-like systems, the production of gel forming recombinant gelatin and secreted production of triple helical or fibre-forming collagen are challenges that still need to be addressed. If realized, they are likely to lead to the development of new products and applications.

1.7 AIM AND THESIS OUTLINE

Animal-derived collagen and gelatin have been extensively used in the past decades for several pharmaceutical and biomedical applications. However, there is need for collagen-based materials with predictable and tailorable properties.

The aim of this thesis is the design and microbial production of gel forming non-hydroxylated collagen-like proteins. Recombinant protein expression and protein engineering are used to develop collagen-like polymers with defined composition, structure, and tunable physical-chemical properties. The possibility of using these proteins as controlled release systems is also explored, as well as the set-up of efficient and scalable production procedures using *P. pastoris* as a microbial factory.

In **Chapter 2**, the design of non-hydroxylated gel-forming collagen-inspired triblock copolymers is described. The production in yeast and preliminary characterization of two such protein polymers is reported.

In **Chapter 3,** four versions of collagen-inspired triblock copolymers, differing only in their mid-block length or amino acid sequence, are used to study the relationship between mid-block size and hydrogel-forming properties. An analytical model, based on classical gel theory is used to help interpreting the experimental data.

In **Chapter 4**, the possibility of using the hydrogels from collagen-inspired triblock copolymers as controlled release systems is investigated. The erosion and protein release kinetics of two hydrogel-forming collagen-inspired triblock copolymers, differing only in mid-block length, is studied and compared.

In **Chapter 5**, the development of a pilot-scale strategy for the fermentation and purification of a new class of gel-forming collagen-like proteins is discussed. Five different collagen-inspired triblock co-polymers are fermented and purified. The pilot-scale production of each of these proteins was initiated to foster high productivity of protein for future applications and research.

In **chapter 6**, some of the results obtained in the thesis are highlighted and suggestions for further research are given.

REFERENCES

1. K. Sanford and M. Kumar, *Curr Opin Biotechnol*, 2005, **16**, 416-421.
2. R. Dandu, A. Von Cresce, R. Briber, P. Dowell, J. Cappello and H. Ghandehari, *Polymer*, 2009, **50**, 366-374.
3. H. Heslot, *Biochimie*, 1998, **80**, 19-31.
4. D. N. Woolfson and M. G. Ryadnov, *Curr Opin Chem Biol*, 2006, **10**, 559-567.
5. S. Kyle, A. Aggeli, E. Ingham and M. J. McPherson, *Trends Biotechnol*, 2009, **27**, 423-433.
6. W. Y. Aalbersberg, R. J. Hamer, P. Jasperse, H. H. J. Jongh, C. G. Kruif, P. Walstra and F. A. De Wolf, *Industrial proteins in perspective*, Elsevier, Amsterdam, 2003.
7. J. Baez, D. Olsen and J. W. Polarek, *Appl Microbiol Biotechnol*, 2005, **69**, 245-252.
8. G. Morreale, E. G. Lee, D. B. Jones and A. P. Middelberg, *Biotechnol Bioeng*, 2004, **87**, 912-923.
9. L. Stryer, *Biochemistry*, 3rd ed. edn., Freeman, New York, 1988.
10. J. Myllyharju and K. I. Kivirikko, *Ann Med*, 2001, **33**, 7-21.
11. D. J. Prockop and K. I. Kivirikko, *Annu Rev Biochem*, 1995, **64**, 403-434.
12. G. Veit, B. Kobbe, D. R. Keene, M. Paulsson, M. Koch and R. Wagener, *J. Biol. Chem.*, 2006, **281**, 3494-3504.
13. J. A. M. Ramshaw, Y. Y. Peng, V. Glattauer and J. A. Werkmeister, *J Mater Sci: Mater Med*, 2008, **20**, 3-8.
14. D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen and J. Polarek, *Adv Drug Deliv Rev*, 2003, **55**, 1547-1567.
15. W. Friess, *Eur. J. Pharm. Biopharm.*, 1998, **45**, 113-136.
16. C. Yang, P. J. Hillas, J. A. Baez, M. Nokelainen, J. Balan, J. Tang, R. Spiro and J. W. Polarek, *BioDrugs*, 2004, **18**, 103-119.
17. S. Young, M. Wong, Y. Tabata and A. G. Mikos, *J Control Release*, 2005, **109**, 256-274.
18. M. Sutter, J. Siepmann, W. E. Hennink and W. Jiskoot, *J Control Release*, 2007, **119**, 301-312.
19. R. Rohanizadeh, M. V. Swain and R. S. Mason, *J Mater Sci Mater Med*, 2008, **19**, 1173-1182.
20. C. H. Lee, A. Singla and Y. Lee, *Int J Pharm*, 2001, **221**, 1-22.
21. Y. Maeshima, P. C. Colorado, A. Torre, K. A. Holthaus, J. A. Grunkemeyer, M. B. Ericksen, H. Hopfer, Y. Xiao, I. E. Stillman and R. Kalluri, *J Biol Chem*, 2000, **275**, 21340-21348.
22. G. M. Cooper, *The Cell: a molecular approach*, 2nd ed. edn., ASM Press, Washington, D. C., 2000.
23. G. N. Ramachandran and A. H. Reddi, *Biochemistry of collagen*, Plenum Press, New York.
24. B. Brodsky, G. Thiagarajan, B. Madhan and K. Kar, *Biopolymers*, 2008, **89**, 345-353.
25. K. H. Gustavson, *Nature*, 1955, **175**, 70-74.
26. R. E. Burge and R. D. Hynes, *Nature*, 1959, **184**, 1562-1563.
27. K. A. Piez and J. Gross, *J Biol Chem*, 1960, **235**, 995-998.
28. S. Sakakibara, Y. Kishida, Y. Kikucha, R. Sakai and K. Kikiuchi, *Bull. Chem. Soc. Jap.*, 1968, **41**, 1273-1275.
29. Y. Kobayashi, R. Sakai, K. Kakiuchi and T. Isemura, *Biopolymers*, 1970, **9**, 415-425.
30. S. Sakakibara, K. Inouye, K. Shudo, Y. Kishida, Y. Kobayashi and D. J. Prockop, *Biochim Biophys Acta*, 1973, **303**, 198-202.
31. P. D. Toman, G. Chisholm, H. McMullin, L. M. Giere, D. R. Olsen, R. J. Kovach, S. D. Leigh, B. E. Fong, R. Chang, G. A. Daniels, R. A. Berg and R. A. Hitzeman, *J Biol Chem*, 2000, **275**, 23303-23309.
32. D. Olsen, J. Jiang, R. Chang, R. Duffy, M. Sakaguchi, S. Leigh, R. Lundgard, J. Ju, F. Buschman, V. Truong-Le, B. Pham and J. W. Polarek, *Protein Expr Purif*, 2005, **40**, 346-357.

33. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
34. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
35. J. Engel and G. Schwarz, *Angew Chem Int Ed Engl*, 1970, **9**, 389-400.
36. A. V. Persikov, J. A. Ramshaw and B. Brodsky, *Biopolymers*, 2000, **55**, 436-450.
37. A. V. Persikov, J. A. Ramshaw, A. Kirkpatrick and B. Brodsky, *Biochemistry*, 2000, **39**, 14960-14967.
38. A. V. Persikov, J. A. Ramshaw, A. Kirkpatrick and B. Brodsky, *J Mol Biol*, 2002, **316**, 385-394.
39. W. Yang, V. C. Chan, A. Kirkpatrick, J. A. Ramshaw and B. Brodsky, *J Biol Chem*, 1997, **272**, 28837-28840.
40. J. Myllyharju, *Matrix Biol*, 2003, **22**, 15-24.
41. F. A. de Wolf and R. C. A. Keller, *Progr. Colloid. Polym. Sci.*, 1996, **102**, 9-14.
42. K. Kuehn, J. Engel, B. Zimmermann and W. Grassmann, *Arch Biochem Biophys*, 1964, **105**, 387-403.
43. J. M. Saddler and P. J. Horsey, *Anaesthesia*, 1987, **42**, 998-1004.
44. L. Ala-Kokko, J. Hyland, C. Smith, K. I. Kivirikko, S. A. Jimenez and D. J. Prockop, *J Biol Chem*, 1991, **266**, 14175-14178.
45. A. S. Olsen, A. E. Geddis and D. J. Prockop, *J Biol Chem*, 1991, **266**, 1117-1121.
46. M. Tomita, N. Ohkura, M. Ito, T. Kato, P. M. Royce and T. Kitajima, *Biochem J*, 1995, **312(Pt 3)**, 847-853.
47. T. Pihlajamaa, M. Perala, M. M. Vuoristo, M. Nokelainen, M. Bodo, T. Schulthess, E. Vuorio, R. Timpl, J. Engel and L. Ala-Kokko, *J Biol Chem*, 1999, **274**, 22464-22468.
48. D. C. John, R. Watson, A. J. Kind, A. R. Scott, K. E. Kadler and N. J. Bulleid, *Nat Biotechnol*, 1999, **17**, 385-389.
49. M. Tomita, H. Munetsuna, T. Sato, T. Adachi, R. Hino, M. Hayashi, K. Shimizu, N. Nakamura, T. Tamura and K. Yoshizato, *Nat Biotechnol*, 2003, **21**, 52-56.
50. F. Ruggiero, J. Y. Exposito, P. Bournat, V. Gruber, S. Perret, J. Comte, B. Olagnier, R. Garrone and M. Theisen, *FEBS Lett*, 2000, **469**, 132-136.
51. C. Merle, S. Perret, T. Lacour, V. Jonval, S. Hudaverdian, R. Garrone, F. Ruggiero and M. Theisen, *FEBS Lett*, 2002, **515**, 114-118.
52. D. D. Buechter, D. N. Paolella, B. S. Leslie, M. S. Brown, K. A. Mehos and E. A. Gruskin, *J Biol Chem*, 2003, **278**, 645-650.
53. I. Goldberg, A. J. Salerno, T. Patterson and J. I. Williams, *Gene*, 1989, **80**, 305-314.
54. P. Li, A. Anumanthan, X. G. Gao, K. Ilangovan, V. V. Suzara, N. Duzgunes and V. Renugopalakrishnan, *Appl Biochem Biotechnol*, 2007, **142**, 105-124.
55. J. L. Cereghino and J. M. Cregg, *FEMS Microbiol Rev*, 2000, **24**, 45-66.
56. I. Keizer-Gunnink, A. Vuorela, J. Myllyharju, T. Pihlajaniemi, K. I. Kivirikko and M. Veenhuis, *Matrix Biol*, 2000, **19**, 29-36.
57. M. Nokelainen, H. Tu, A. Vuorela, H. Notbohm, K. I. Kivirikko and J. Myllyharju, *Yeast*, 2001, **18**, 797-806.
58. E. C. de Bruin, F. A. de Wolf and N. C. Laane, *Enzyme Microb Technol*, 2000, **26**, 640-644.
59. P. Bruckner and D. J. Prockop, *Anal Biochem*, 1981, **110**, 360-368.
60. J. M. Davis and H. P. Bachinger, *J Biol Chem*, 1993, **268**, 25965-25972.
61. E. C. de Bruin, M. W. Werten, C. Laane and F. A. de Wolf, *FEMS Yeast Res*, 2002, **1**, 291-298.
62. O. Pakkanen, E. R. Hamalainen, K. I. Kivirikko and J. Myllyharju, *J Biol Chem*, 2003, **278**, 32478-32483.

63. D. R. Olsen, S. D. Leigh, R. Chang, H. McMullin, W. Ong, E. Tai, G. Chisholm, D. E. Birk, R. A. Berg, R. A. Hitzeman and P. D. Toman, *J Biol Chem*, 2001, **276**, 24038-24043.
64. B. Montagnon, J. C. Vincent-Falquet and B. Fanget, *Dev. Biol. Stand.* , 1984, **55**, 37– 38.

Chapter 2

Precision gels from collagen-inspired triblock copolymers

This chapter as been published in modified form as: M. W. T. Werten, H. Teles, A. P. H. A. Moers, E. J. H. Wolbert, J. Sprakel, G. Eggink, and F. A. de Wolf, *Biomacromolecules*, 2009, **10**, 1106-1113.

SUMMARY

Gelatin hydrogels find broad medical application. The current materials, however, are from animal sources, and their molecular structure and thermal properties cannot be controlled. This study describes recombinant gelatin-like polymers with a general design that inherently offers independent tuning of the cross-link density, melting temperature, and biocompatibility of the gel. The polymers contain small blocks with thermoreversible trimerization capacity and defined melting temperature, separated by hydrophilic nontrimerising blocks defining the distance between the knot-forming domains. As an example, we report the secreted production in yeast at several g/l of two non-hydroxylated ~42 kDa triblock copolymers with terminal trimerising blocks. Because only the end-blocks formed cross-links, the molecular architecture of the gels is much more defined than that of traditional gelatins. The novel hydrogels had a ~37 °C melting temperature, and the dynamic elasticity was independent of the thermal history. The concept allows to produce custom-made precision gels for biomedical applications.

2.1 INTRODUCTION

Stimulus-responsive polymer gels are promising materials for chemomechanical systems, sensors, surgery, regenerative medicine, and pharmaceutics[1]. Biocompatibility is an essential requirement for medical gels[2] and, therefore, gelatin[3], which is derived from the natural matrix protein collagen[4], is ideally suited. To the benefit of millions of patients, several kilotons[5] of gelatin with a billion dollar market value are used on an annual basis and novel gelatin materials are being developed for bone[6], retinal[7], tissue-enhancing[8], and other implants, in vascular prostheses[9], sponge embolization therapy[10], scaffolds for tissue and cell culture[11, 12], medical glues[13], blood supplementation fluids[14, 15], drug carriers[11, 16], wound dressings[17], and vaccines[18].

For such applications, the animal origin of gelatin poses the risk of contamination with transmissible disease agents (e.g., prions) and immune responses[15, 18-20.] Furthermore, the extraction of gelatin from bones and hides causes uncontrolled degradation and results in a variable multitude of molecular species[3]. Recombinant production allows to circumvent these problems and, accordingly, several groups including our own have produced single-component natural gelatins[5, 20, 21] and non-natural gelatin-like designer proteins[22-25]. We realised secreted production of gelatins in the yeast *Pichia pastoris* at very high levels, as well as simple and scalable purification procedures[21, 25]. In contrast to animal gelatins, these recombinant gelatins are naturally nongelling.

Nongelling recombinant gelatins are advantageous when used in blood supplementation fluids and vaccines, where the current use of animal gelatins requires suppression of unwanted gel formation by chemical modification[14, 15, 26] or degradation[18, 20], respectively. However, most medical applications of gelatin do require gel formation. The basis of the thermoreversible knots in a gelatin gel is the characteristic triple-helical structure of collagen[3]. It consists of repetitive Gly-Xaa-Yaa triplet repeats, commonly containing approximately 22 % proline[3]. At this low proline content, triple helices are not thermally stable above 5-15 °C unless, as occurs in animals, the prolines in the Yaa position are post-translationally modified to 4-hydroxyprolines by the enzyme peptidyl-prolyl-4-hydroxylase (P4H)[27-29]. Because microbial hosts generally lack this enzyme, recombinant production of thermally stable triple-helical collagen and gelatin requires coexpression of the two subunits of the mammalian P4H[27, 28, 30].

For applications such as controlled drug delivery, independent control over the melting temperature (T_m), cross-link density, and biocompatible properties of the gelatin gel would be highly advantageous. Thus far, however, no such material has been conceived. Manipulating the T_m of recombinant gelatins is already difficult in itself, as it would require control over the degree of hydroxylation. Proline-rich trimer-forming gelatin-like (collagenous) peptides with adjustable, length-dependent T_m have been synthesised chemically[31-33]. but these are too short to form gels and cannot be produced in an economically feasible way. Polycondensated (and, thus, polydisperse) peptides long enough to form a gel can neither provide a homogeneous, controllable, and biomedically relevant T_m, nor a controllable cross-link density.

We describe here recombinant non-hydroxylated gelatin-like protein polymers capable of forming tuned precision gels. The block copolymers consist of short proline-rich triple helix-forming blocks that define the T_m, separated by long random coil spacer blocks that define the distance between the knot-forming domains in the molecules. The currently used spacers are extremely hydrophilic[25]. At least one of them is capable of attracting human cells in culture[34], and has favorable biocompatibility as compared to animal products in blood applications[15]. Gel-forming triblock protein polymers have been described before, making use of coiled-coils[35-40], elastin-mimetic motifs[41], or the fluorescent protein DsRed[39] as the cross-link-forming domain. The present work is the first report of such polymers where the knots are formed by (trimerising) collagenous blocks. This choice of cross-linking domain offers thermoreversible melting and a defined junction multiplicity[42] (number of associative groups per knot) of exactly three. Notably, also the random coil spacer blocks are of a collagenous (gelatin-like) nature. The polymers, thus, truly represent a novel class of gelatins, offering defined functionality (number of associative groups per molecule) and tunable T_m, while expectedly retaining the hallmark biocompatibility of traditional collagenous materials.

2.2 MATERIAL AND METHODS

Construction of expression vectors

Triple helix-forming block T was prepared by PCR using the oligonucleotides T-FW and T-RV (apendix 2, Table 2.2.1). The ~0.1 kb product was cloned into vector pCR4-TOPO (Invitrogen), resulting in vector pCR4-TOPO-T. The previously described vector pMTL23-P4[25] contains a gene encoding the custom-designed, highly hydrophilic 36.8 kDa collagenous protein "P4". This vector was digested with *Dra*III (5' to the P4 gene) and dephosphorylated. Vector pCR4-TOPO-T was digested with *Dra*III/*Van*91I. The released T block was ligated into the linearised and dephosporylated vector, to yield vector pMTL23-TP4. This vector was then digested with *Van*91I (3' to the P4 gene) and dephosphorylated, and a second *Dra*III/*Van*91I digested T block was inserted to yield vector pMTL23-TP4T. The TP4T gene was cloned into *P. pastoris* expression vector pPIC9 (Invitrogen) via *Xho*I/*Eco*RI.

The gene encoding R4 was constructed by concatenating four copies of an R gene monomer. The monomeric gene was designed by randomising the sequence of the P gene monomer[25] in such a way that not every third residue of the encoded protein is glycine, preventing the formation of collagen triple helices while maintaining the same amino acid composition. The R gene monomer was constructed by overlap extension PCR[43] using oligonucleotides RA-FW and RA-RV for the 5' half of the gene, and oligonucleotides RB-FW and RB-RV for the 3' half (apendix 2, Table 2.2.1). The products of these reactions were combined by overlap extension PCR to generate the entire R gene monomer. The monomer was cloned into vector pMTL23[44] via

*Xho*I/*Eco*RI, and multimerised via *Dra*III/*Van*91I as described previously for P4[25], resulting in vector pMTL23-R4. The R4 fragment was cloned into expression vector pPIC9 via *Xho*I/*Eco*RI. The construction of pMTL23-TR4T and subsequent subcloning into pPIC9 was analogous to the procedures described for TP4T.

The DNA sequences (and translated amino acid sequences) of the *Xho*I/*Eco*RI fragments encoding P4[25], R4, TP4T, and TR4T have been deposited in GenBank under accession numbers EU834225-EU834228. Although, for clarity, the T blocks are referred to as $(\text{Pro-Gly-Pro})_9$ in the main text, the cloning procedure results in a Gly-Pro-Pro-Gly-Ala extension at the N-terminus, and an Ala-Gly-Gly extension at the C-terminus.

Transformation of *P. pastoris*

Expression vectors were linearised with *Sal*I to promote integration at the *his4* locus rather than the *AOX1* locus, thus enabling normal growth on methanol[45]. Transformation of *P. pastoris his4* strain GS115[45] and selection of transformants was as described previously[21].

Fermentation of *P. pastoris*

Fed-batch fermentations were performed in 2.5-L Bioflo 3000 fermenters (New Brunswick Scientific), essentially as described by Zhang et al.[46]. Minimal basal salts medium[46] was used and no protease-inhibiting supplements were added. The pH was maintained at 3.0 throughout the fermentation by addition of ammonium hydroxide as base. The methanol fed-batch phase for protein production lasted two to three days. A homemade semiconductor gas sensor-controller, similar to that described by Katakura et al.[47], was used to monitor the methanol level in the off-gas and to maintain a constant level of ~0.2 % (w/v) methanol in the broth. At the end of the fermentation, the cells were separated from the broth by centrifugation for 10 min at 10000 × g (RT) in an SLA-3000 rotor (Sorvall), and the supernatant was microfiltered.

Protein purification

All centrifugation steps were performed in an SLA-1500 or SLA-3000 rotor (Sorvall) for 30 min at 20000 × g, and resuspension of protein pellets was always in Milli-Q water at 65 °C.

As a precaution, the cellfree broth was heated for 30 min at 65 °C to melt possible gel structures formed by the recombinant protein. The pH was raised to 8.0 by addition of sodium hydroxide to allow precipitation of medium salts by centrifugation (RT). The protein was precipitated from the supernatant by addition of ammonium sulfate to 40 % of saturation, followed by incubation on ice for 30 min and centrifugation (4 °C). This precipitation procedure was repeated once, and 50 mM sodium chloride and 40 % (v/v) acetone were added to the final resuspended pellet. After centrifugation (4 °C), the pellet was discarded and acetone was added to the supernatant up to 80 % (v/v). The protein pellet obtained after centrifugation was air-dried, resuspended, and desalted by extensive dialysis against Milli-Q water. The final product was lyophilised.

SDS-PAGE and N-terminal protein sequencing

The NuPAGE Novex system (Invitrogen) was used for SDS-PAGE, with 10 % Bis-Tris gels, MES SDS running buffer, and SeeBlue Plus2 prestained molecular mass markers. Gels were stained with Coomassie SimplyBlue SafeStain (Invitrogen). Blotting of proteins for N-terminal sequencing by Edman degradation was as described previously[21]. Protein sequencing was performed by Midwest Analytical (St. Louis, MO).

Mass spectrometry

Matrix-assisted laser desorption/ionisation (MALDI) mass spectrometry was performed using an Ultraflex mass spectrometer (Bruker). Samples were prepared by the dried droplet method on a 600 µm AnchorChip target (Bruker), using 5 mg/ml 2,5-dihydroxyacetophenone, 1.5 mg/ml diammonium hydrogen citrate, 25 % (v/v) ethanol, and 1 % (v/v) trifluoroacetic acid as matrix. Measurements (50 shots at 20 Hz) were made in the positive, linear mode, with the following parameters: ion source 1, 20000 V; ion source 2, 18450 V; lens, 5000 V; pulsed ion extraction, 550 ns. Protein Calibration Standard II (Bruker) was used for external calibration.

Circular dichroism spectroscopy

Proteins were dissolved at 100 mg/ml in 0.2 M sodium phosphate pH 3.0 and heated to 45 °C. Gels were allowed to form overnight inside the cuvette (path length 0.01 mm) at room temperature. Circular dichroism (CD) spectra were recorded from 260 to 210 nm using a J-715 spectropolarimeter (Jasco) set at the desired temperature. Spectra were obtained as the average of ten consecutive scans using a scanning speed of 100 nm/min at a resolution of 0.2 nm. Melting curves from 20 to 65 °C were derived by increasing the temperature in steps of 5 or 10 °C, followed by equilibration at each temperature for 10 min prior to scanning. To allow recording of full spectra from 260 to 195 nm, measurements were also performed at 0.2 mg/ml in a cuvette with 1 mm path length.

Differential scanning calorimetry

Degassed 0.5 ml protein solutions (0.4-2.2 mM in 0.2 M sodium phosphate pH 3.0) were loaded into a MicroCal VP-DSC calorimeter at 45 °C. Protein solutions were equilibrated for 10 h at 20 °C to allow complete triple helix formation. For melting curves, the temperature was increased from 5 to 65 °C at a scan rate of 15 °C/h.

Rheometry

An Anton Paar Physica MCR301 rheometer, equipped with a stainless steel CC17 Couette geometry, gap size, 0.71 mm; bob radius, 8.3 mm; and sample volume 3 ml, was operated at an angular frequency of 1 Hz and a strain of 0.1 %. Protein solutions (2.2 mM in 0.2 M sodium phosphate, pH 3.0) were heated to 45 °C, introduced in the geometry, and subsequently quenched to 20 °C to induce gel formation. Measurements were done at 20 and 5 °C. Melting of the gel was studied by increasing the temperature from 20 to 65 °C in steps of 2.5 °C that lasted 10 min each, and these data were smoothed using a moving average filter over 21 data points (10 min).

Scanning electron microscopy

TR4T was dissolved at 100 mg/ml in Milli-Q water at 60 °C. A gel was allowed to form overnight at room temperature and then glued onto a brass sample holder using Tissue-Tek (TBS). The specimen was freeze-fractured in melting nitrogen slush, vacuum transferred to a CT1500 cryostage (Oxford Instruments), and sputter-coated with ~10 nm of platinum. After cryoetching for ~5 min the surfaces were analyzed at -184 °C using a Field-Emission Scanning Electron Microscope (JEOL 6300 F) at a working distance of ~10 mm, with secondary electron detection at 3.5 kV. Images were recorded digitally using Orion 6 (ELI) at a scan rate of 100 s (full frame) at a size of 2528 x 2030 pixels, 8 bit. Image noise reduction was done in Orion 6, and saturation, contrast, and brightness were adjusted in Adobe Photoshop CS2.

Erosion and selease studies

The erosion rate of TR4T gels was measured using a method similar to that reported previously[38]. TR4T was dissolved at 200 mg/ml in PBS (10 mM sodium phosphate, 150 mM sodium chloride, pH 7.4) at 60 °C. Gel layers of 1 mm thickness were made by pipetting 44 µL of the solution into cylindrical polypropylene containers of 7.5 mm diameter and 5 mm height. The containers were sealed in plastic foil to prevent drying and the gels were allowed to set overnight at room temperature. The containers were unpacked and submerged in 2.2 ml of PBS and incubated at either 20 or 37 °C under gentle tilting at 20 rpm with an amplitude of ~6 mm. Erosion of TR4T was determined by monitoring the protein concentration in the supernatant by measuring the absorbance at 230 nm (for lack of absorbance at 280 nm owing to the absence of aromatic amino acids in TR4T) and comparison with a calibration curve of TR4T. Enzyme release from TR4T gels was studied similarly, using gels where 10 mg/ml lysozyme was added to the molten gel prior to preparing the gel layers. The release of lysozyme was quantified by determining the tryptophan fluorescence (Ex = 295 nm, Em = 240 nm) of the supernatant over time against a calibration curve of lysozyme (eroded TR4T did not cause significant background fluorescence as it does not contain aromatic residues). Both the erosion and the release experiments were performed in triplicate, and absorbance and fluorescence measurements were done using a Safire microplate reader (Tecan).

2.3 RESULTS

Rationale for non-hydroxylated precision gels

As an example of the proposed concept, two ABA type triblock copolymers designated TP4T and TR4T are described, but many other functional multiblock configurations are conceivable. The amino acid sequences of all protein polymers described in this work are available through GenBank (accession numbers ACF33476-ACF33479).

The triple helix-forming T blocks at both ends of the triblock copolymers consist of $(\text{Pro-Gly-Pro})_n$ homopolymeric stretches (Figure 2.1a). On the basis of helix melting studies with chemically synthesised $(\text{Pro-Pro-Gly})_n$ [32, 33, 48], the length of n was tentatively chosen to be nine to provide a melting point in a biomedically relevant range.

The middle section of the molecule acts as a hydrophilic spacer (Figure 2.1a) and consists of either of two random coiled block variants. The first is P4, a synthetic gelatin-like molecule previously developed by our group[25]. It is extremely hydrophilic, acts as a cytophilic protein in human cell culture[34], and shows outstanding biocompatibility as a plasma expander[15]. Despite P4's collagenous primary structure, it does not form detectable triple helices at 4 °C because of the absence of 4-hydroxyprolines[25]. It cannot, however, be excluded beforehand that P4 might play a minor role in network formation in the presence of the proline-rich trimer-forming T blocks. Therefore, a second type of mid-block, R4, was constructed. It has the same amino acid composition as P4, but its protein sequence is quasi-random in that it does not have glycine as every third residue. Thus, R4 by definition cannot form triple helices.

Because both the end-blocks and mid-blocks have a defined length and composition, and because trimerisation of the mid-blocks is unlikely (P4) or impossible (R4), the resulting gels are expected to contain only cross-links made-up of (Pro-Gly-Pro)$_9$ (Figure 2.1b-c). This is in stark contrast with traditional gels prepared from animal gelatin, where all molecules have a different makeup, and the entire chain engages in network formation.

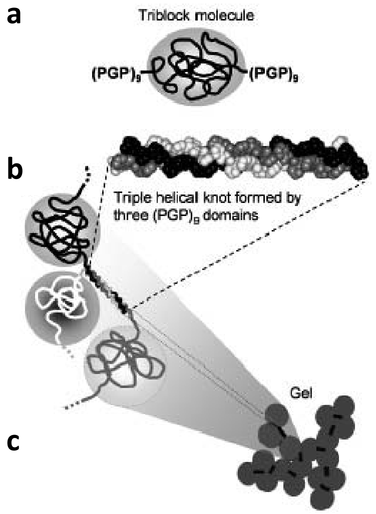

Figure 2.1 Schematic representation of gel formation by triblock copolymers. (a) Triblock copolymer with (Pro-Gly-Pro)$_9$ end-blocks and random coil mid-block. (b) Detail of a (Pro-Gly-Pro)$_9$ end-block in triple-helical form and of a trimeric knot in the gel. Three mid-blocks originate from the knot, each of which also has a (Pro-Gly-Pro)$_9$ block at its other end (not shown for clarity). (c) Gel consisting of triblock copolymers. The circles represent the time-averaged space occupied by the random coiled mid-blocks. The dark bars represent a random subset of trimeric knots formed between three neighboring chains.

Biosynthesis of triblock copolymers

P. pastoris strains expressing TR4T and TP4T as extracellular proteins were constructed and grown in bioreactors. Culture supernatants were analyzed by SDS-PAGE (Figure 2.2a). The proteins were purified from the cell-free broth essentially by differential ammonium sulfate precipitation, similarly to the purification of P4[25]. The purity of the proteins was estimated to be at least 99 %, based on amino acid analysis and subsequent linear least-squares fitting to the observed data of (1) the theoretical composition of the respective pure protein and (2) the composition determined for host-derived proteins present in the medium. Purified TR4T and TP4T migrated as single bands in SDS-PAGE (Figure 2.2a), indicating their purity and intactness. MALDI mass spectrometry confirms this observation and shows that the molecular weight of the proteins is within experimental error of the expected value of 41741 Da (Figure 2.2b). Clearly, the proteins migrate abberantly in SDS-PAGE, as was demonstrated previously for P4[25]. N-terminal sequencing of the bands separated by SDS-PAGE further confirmed the identity of the products. A minor fraction (~15 %) of the molecules had a single N-terminal Glu-Ala extension, which is known to occasionally occur because of partial processing of the α-factor prepro secretory signal by the *P. pastoris* dipeptidylaminopeptidase[21]. Judging from the intensity of the bands in SDS-PAGE, the volumetric productivity of TR4T and TP4T appears comparable to that of P4, which is produced at 3-6 g/l of clarified broth[25].

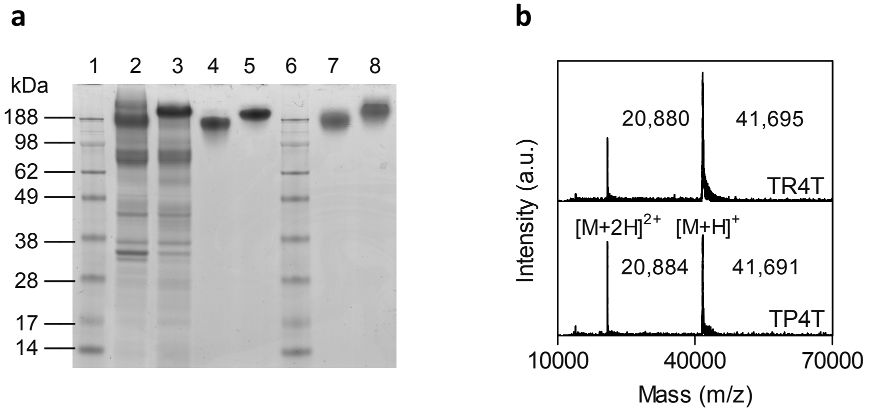

Figure 2.2 Production of protein polymers. (a) SDS-PAGE: Lane 1, 6, molecular weight marker; lane 2, culture supernatant of TR4T; lane 3, culture supernatant of TP4T; lane 4, purified TR4T; lane 5, purified TP4T; lane 7, purified control protein R4; lane 8, purified control protein P4. For culture supernatants, 5 μl was loaded, and for purified proteins ~20 μg. (b) MALDI of purified TR4T and TP4T. Singly and doubly charged molecular ions are indicated.

Trimer-forming functionality is restricted to proline-rich end-blocks

The occurrence of trimeric knots is essential for the generation of a thermoreversible gelatin network. In the triblock design, only the proline-rich T blocks at both ends of the triblock copolymers fulfill this role, and the mid-blocks act as spacers. To confirm that the combined blocks in the triblock copolymer retain these distinct functions, UV-CD spectroscopy was performed at a protein concentration of 100 mg/ml, where the chance of intermolecular interactions is high. Both TR4T and TP4T formed optically clear dispersions and yielded very similar spectra, with a consistently negative ellipticity below 260 nm and a negative peak around 205-200 nm, reflecting mainly random structure[31, 49]. Spectra obtained with TR4T and TP4T at 20 and 65 °C are shown in Figures 2.3a-b (Figures 2.3c-d show measurements at low concentration that permitted registration of spectra down to 195 nm). The predominance of random structure is in agreement with the random coil propensity of the large mid-block, making up 88 % of the molecular length. To investigate whether the secondary structure of free mid- and end-blocks can fully account for that of TR4T and TP4T, or whether the mid- and end-blocks influence each other's structure when coupled, the triblock spectra were compared to those of the individual blocks. Linear least-squares fitting of spectra of isolated R4 (Figure 2.3e) and Ac-(Gly-Pro-Pro)$_{10}$-NH$_2$ in triple-helical or heat-denatured form[50] (data kindly provided by H.P. Bächinger) to the spectra of TR4T resulted in a close fit (Figure 2.3a). A poly-(Gly-Pro-Pro) content of ~9 % was obtained both at 20 and 65 °C, in reasonable agreement with the 12 % length ratio of the end-blocks relative to the entire block copolymer. This shows that the end and mid-blocks in TR4T indeed assume the same secondary structure as the isolated components and, thus, that at 20 °C, the end-blocks, like Ac-(Gly-Pro-Pro)$_{10}$-NH$_2$[50], are largely in helical conformation. The spectral contribution of the end-blocks can be appreciated by subtraction of the fitted contribution of R4 from the spectra of TR4T at 20 and 65 °C (see inset Figure 2.3a). Similarly, the secondary structure of TP4T could be fully accounted for by that of isolated P4 (Figure 2.3f) and Ac-(Gly-Pro-Pro)$_{10}$-NH$_2$ (see inset Figure 2.3b), and its poly-(Gly-Pro-Pro) content of ~14 % closely matched the expected value. These results show that, in both polymers, selectively, the proline-rich end-blocks behave like Ac-(Gly-Pro-Pro)$_{10}$-NH$_2$[50] and, thus, take part in helix formation.

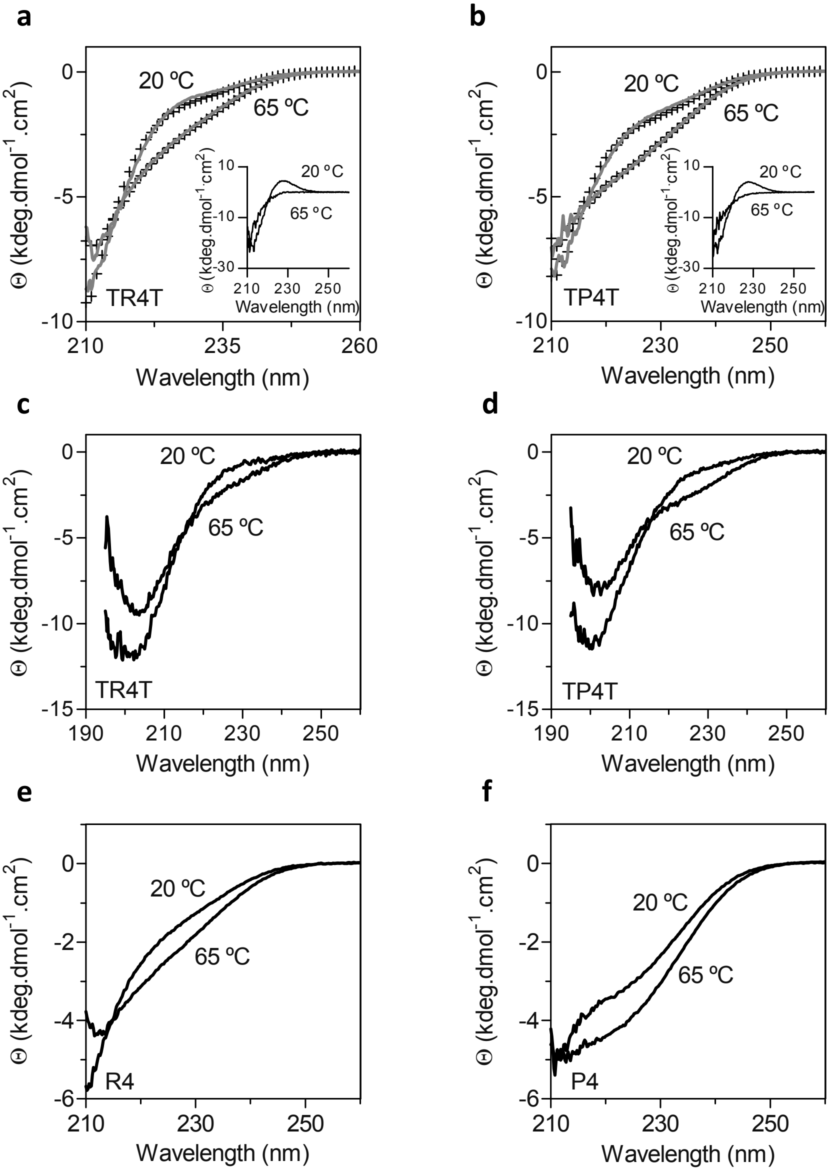

Figure 2.3 Study of the secondary structure at 20 and 65 °C, as reflected by CD. (a) Spectra of TR4T (100 mg/ml) and (b) TP4T (100 mg/ml) are shown, along with fitted spectra (+) composed

of contributions of R4 and P4, respectively, and (Gly-Pro-Pro)$_n$ in isolated form. Inset: spectra of the proline-rich end-blocks, as deduced by subtracting the fitted contribution of R4 and P4 from the spectra of TR4T and TP4T, respectively. (c) Spectra of TR4T (0.2 mg/ml). (d) Spectra of TP4T (0.2 mg/ml). (e) Reference spectra of free R4 (100 mg/ml). (d) Reference spectra of free P4 (100 mg/ml).

Melting behavior of triple-helical end-blocks

The average ellipticity from 230 to 225 nm (at 100 mg/ml), as an indicator of the presence of 3_i helices[49], was determined by circular dichroism spectroscopy at various temperatures between 20 and 65 °C. Figures 2.4a-b show the thermal response of the R4 and P4 (monoblock) polymers, which, in view of the above-described absence of crosstalk between end and mid-blocks, reflect the properties of the mid-block in TR4T and TP4T, respectively. Its near-linear temperature dependence is characteristic of a gradual repopulation of local conformations of individual single chain sections[49, 51, 52]. The behavior of the end-blocks can be deduced by subtracting the contribution of the free R4 and P4 monoblock polymers, as determined by component fitting (Figure 2.3), from the melting curve of TR4T and TP4T, respectively. The observed sigmoidal temperature dependence (Figures 2.4a-b), is characteristic of cooperative unfolding of triple helices[31, 33, 48]. The apparent T_m for the proline-rich end-blocks deduced from the inflection point of these curves was 47 °C for both TR4T and TP4T. The corresponding van ’t Hoff enthalpies were calculated to be 188 and 223 kJ/mol, respectively. Based on the calorimetric enthalpy obtained for (Pro-Pro-Gly)$_{10}$ trimers by Frank et al.[33], the melting enthalpy expected for a (Pro-Gly-Pro)$_9$ end-block trimer is 223 kJ/mol. The van ’t Hoff enthalpies obtained for TR4T and TP4T are 0.85 and 1.0 times this value, respectively. This shows that the cooperatively melting unit is approximately equal to the end-blocks, providing further evidence that triple helix formation occurs specifically through the proline-rich end-blocks.

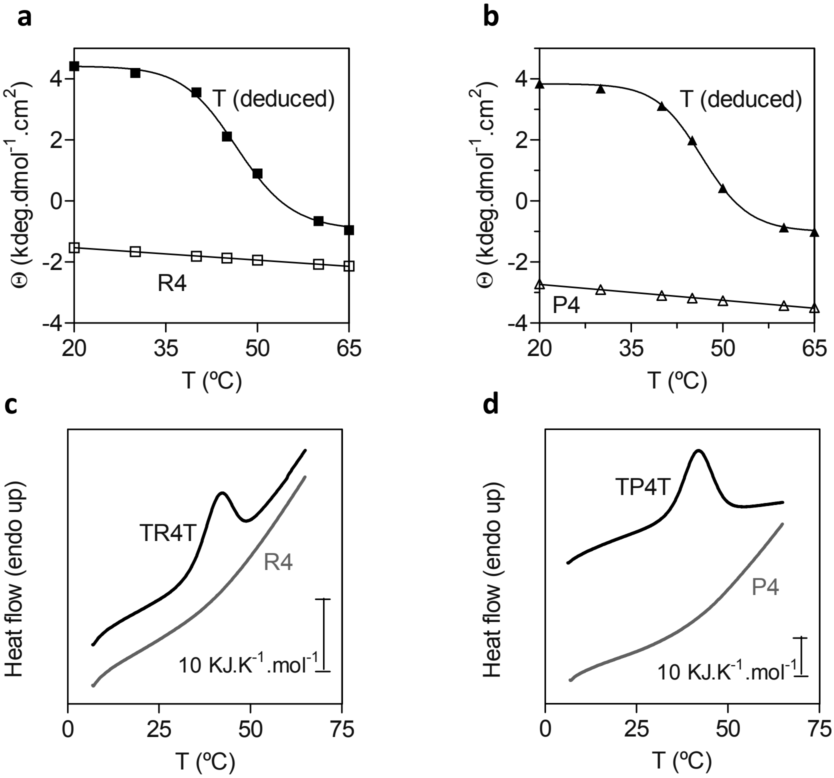

Figure 2.4 Thermal denaturation as reflected by (a) and (b) the average ellipticity from 230 to 225 nm or by (c) and (d) DSC. Temperature profile of (a) R4 and (b) P4 produced as separate monoblock polymers and of the proline-rich end-blocks as deduced by subtracting the fitted contribution of R4 and P4 from the melting curve of TR4T and TP4T, respectively. DSC thermograms of (c) R4 (1.4 mM) and TR4T (1.1 mM) and (d) P4 (1.4 mM) and TP4T (1.1 mM), recorded after equilibration for 10 h at 20 °C.

The thermal denaturation of the polymers between 5 and 65 °C was also studied by differential scanning calorimetry (DSC; Figures 2.4c-d). While R4 and P4 polymer solutions did not show endothermic transitions, and thus, no triple helices were present, TR4T and TP4T had ΔC_p maxima at 41 ± 0.5 °C (±s.d., n = 3) and 42 ± 0.4 °C (±s.d., n = 3), respectively. This confirms the conclusion drawn from CD that selectively the end-blocks in the triblock copolymers are responsible for trimerisation. The somewhat higher T_m values obtained with CD probably result from a systematic offset from the true temperature, owing to inefficient heat transfer in the spring-loaded CD cuvette assembly. The enthalpies of TR4T and TP4T were 154 ± 56 kJ/mol (±s.d., n = 3) and 140 ± 8 kJ/mol (±s.d., n = 3), respectively, in good agreement with the expected enthalpy of about 148 kJ/mol per set of two (Pro-Gly-Pro)$_9$ end-blocks (i.e., two-thirds of the value of an end-block trimer). This indicates once more that nearly all (Pro-Gly-Pro)$_9$ blocks in TR4T and TP4T had formed triple helices. The shape of the thermograms of TR4T and TP4T reveals van ’t Hoff enthalpies of 260 ± 39 kJ/mol (±s.d., n = 3) and 321 ± 79 kJ/mol (±s.d., n = 3), respectively. This corresponds to, respectively, 1.2 ± 0.2 (±s.d., n = 3) and 1.4 ± 0.4 (±s.d., n = 3) times the calorimetric enthalpy of a (Pro-Gly-Pro)$_9$ trimer. Therefore, as was concluded from CD, the cooperatively melting unit roughly corresponds to the end-blocks. A thermal denaturation cycle repeated on the same samples (after 10 h at 20 °C) resulted in calorimetric and van ’t Hoff enthalpies that deviated less than 3 and 6 %, respectively, from the above values obtained with freshly prepared solutions, showing complete reversibility of trimer formation.

Triblock copolymers form gels

Initial observations showed that TR4T and TP4T, unlike R4 and P4[25], formed optically clear gels when preheated solutions were cooled to room temperature. For example, at 7.8 mM equiv of trimer-forming end-block, a stainless steel bead with a diameter of 4.8 mm and a weight of 0.45 g, was supported by a TR4T gel during at least several hours until the gel was molten by heating.

Gel formation was further investigated by means of dynamic rheology, at a protein concentration of 2.2 mM, corresponding to 4.4 mM or ~10 mg/ml equiv of (Pro-Gly-Pro)$_9$ end-blocks (Figure 2.5). Upon quenching from 45 to 20 °C, the storage modulus (dynamic elasticity) steeply increased within minutes, while the loss modulus (dynamic loss) remained low. After roughly five hours at 20 °C, the storage modulus of TR4T and TP4T reached a final value of approximately 380 and 190 Pa, respectively. This range is in good agreement with values obtained in a variety of animal gelatins at a corresponding triple helix concentration[53], which confirms that most of the end-blocks took part in helix-forming knots. No further changes of the storage modulus were observed during the 28-hour measurement. The gel strength increased 5 (TR4T) to 25 % (TP4T) upon cooling from 20 to 5 °C.

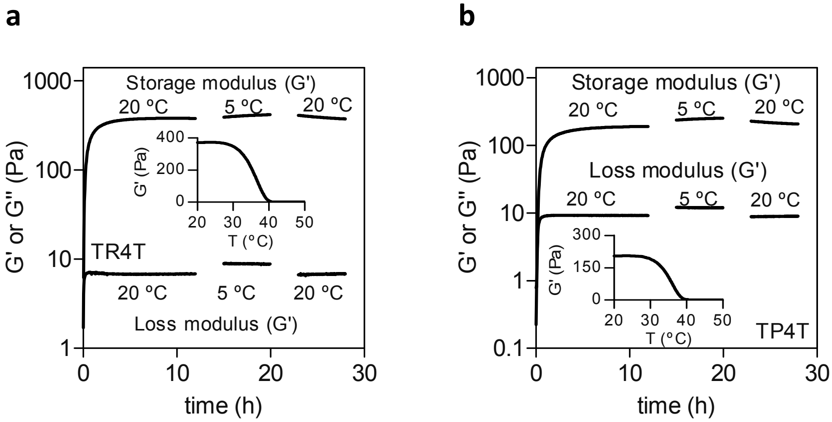

Figure 2.5 Characterisation of (a) TR4T and (b) TP4T gel by dynamic rheology. The time-resolved storage and loss modulus were measured at the temperatures indicated, at a protein concentration of 2.2 mM. For each new temperature, the sample was allowed to equilibrate for three hours in the geometry before monitoring. Insets: storage modulus as a function of temperature.

The melting behavior of the TR4T and TP4T gels was studied by increasing the temperature from 20 to 65 °C and following the decrease of the storage modulus (see insets in Figures 2.5a-b). The apparent inflection point in the melting curves of both gels is at ~37 °C, which is below the T_m values obtained with DSC. This is in agreement with the expectation that melting of the gels occurs before the majority of the trimeric knots unfold.

Hydrogel morphology

As the microstructure of the gels could influence their release characteristics, scanning electron microscopy (SEM) was performed on a freeze-fractured TR4T gel.

The distribution of micropores is highly homogeneous, and an interconnected network structure with smooth interfaces is apparent (Figure 2.6).

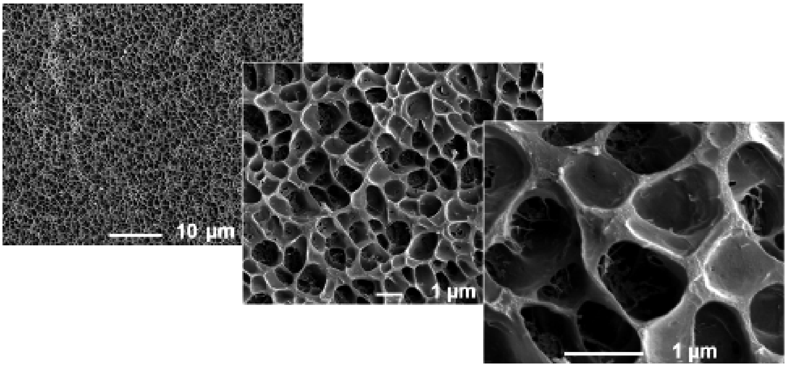

Figure 2.6 SEM micrographs of a freeze-fractured **TR4T** hydrogel (100 mg/ml) at different magnifications.

Enzyme release and hydrogel erosion

To get an impression of the suitability of the present gels for release of, for example, protein drugs, the release from a TR4T gel (200 mg/ml) of the 21 kDa model protein lysozyme, and the erosion of the gel itself (both in a 50-fold excess volume of PBS buffer) were investigated in a pilot study (Figure 2.7). At 20 °C, the release of lysozyme was higher than the erosion of the hydrogel, indicating that release of the enzyme was caused both by erosion of the gel and by diffusion of the enzyme through the polymer network. Assuming erosion at the surface of the gel, the erosion rate at this temperature was 4.6×10^{-3} mg cm^{-2} min^{-1}, as calculated from the slope of the linear regression line indicated in Figure 2.7. At physiological temperature the release of lysozyme was virtually equal to the erosion of the gel. This bulk erosion concurs with the observation that TR4T gels (200 mg/ml) at this temperature exist in a semimolten state, in accordance with the apparent T_m of ~37 °C as determined by rheology (at 100 mg/ml).

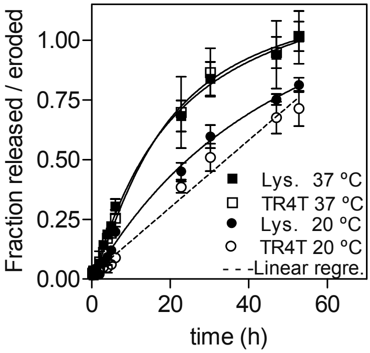

Figure 2.7 Erosion of a TR4T gel (200 mg/ml), and release of lysozyme (10 mg/ml) at 20 and 37 °C. Error bars represent s.d. (n = 3).

2.4 DISCUSSION

Gelatins are important biomedical materials but currently have the disadvantage of being animal-derived, polydisperse, and chemically ill-defined. Furthermore, it is not possible to independently control the T_m, the distance between domains involved in triple-helices (forming the thermoreversible knots in the gel), and the biocompatibility. In this study, we met these challenges through the development and efficient production of recombinant gelatin-like block copolymers.

A prerequisite for the formation of a gelatin network is the formation of triple-helical knots. The proline-rich end-blocks in our triblock design serve this function. The mid-blocks represent random coil spacer regions that cannot readily (P4), or at all (R4), participate in trimer formation. We have shown here that the (Pro-Gly-Pro)$_9$ end-blocks in TP4T and TR4T are both exclusively and near-quantitatively involved in trimerisation, as both the calorimetric enthalpies obtained with DSC and the van't Hoff enthalpies obtained with DSC and CD were in good agreement with values expected for free (Pro-Gly-Pro)$_9$. The cross-links formed by the end-blocks, thus, essentially have a fixed length and amino acid composition and, accordingly, a fixed T_m. Given this concurrence in enthalpic values the alignment of the end-blocks is probably even largely in register, as is known for chemosynthetic (Pro-Pro-Gly)$_{10}$ peptides in trimeric form[54]. In any case, the triblock gelatins are highly defined in that the cross-links are formed strictly by the end-blocks. This is in striking contrast with animal gelatin molecules, which form triple helices along their entire chain. As a consequence, the cross-links in animal gelatin gels have a wide variety of lengths, compositions, and melting temperatures[53]. The fact that traditional gelatins comprise a wide range of different molecules further increases their undefined nature.

Because animal gelatin gels are such highly complex systems, they have very broad melting trajectories. During aging, the storage modulus continues to increase in proportion to the logarithm of time[53, 55]. Furthermore, these gels show a pronounced dependence on the thermal history[53, 55]. For example, Te Nijenhuis[55] found that the stiffness of animal gelatins preaged at 17.4 °C strongly increased upon lowering the temperature, as additional helices of low thermal stability were formed. Our triblock gelatins behave very differently. Only the end-blocks in TP4T and TR4T gels form helices, and these have distinct thermal stability. Consequently, the storage modulus reached a plateau within only a few hours and the stiffness of the gels after preaging at 20 °C hardly increased upon cooling to 5 °C. The T_m of the gels is defined and can be directly controlled by varying the length of the terminal (Pro-Gly-Pro)$_n$ stretches, as known from studies with chemically synthesised peptides[32, 33, 48, 56]. Indeed, preliminary DSC experiments showed that, under identical conditions, triblock copolymers with (Pro-Gly-Pro)$_{16}$ end-blocks have a much higher T_m (~57 °C) than TR4T (~41 °C).

Although the T_m values of TR4T and TP4T may appear rather high as compared to published T_m values for chemosynthetic (Gly-Pro-Pro)$_n$ [32, 48], the T_m is known to depend highly on the peptide concentration[32, 33]. Accordingly, the relatively high concentrations used here are probably the major determinant for the T_m values found. The minor extensions of the end-blocks resulting from the cloning procedure (see Experimental Section), or the possible involvement in trimerisation of a few residues of the mid-blocks (immediately adjacent to the end-blocks) may provide an additional increase in T_m. These effects will, however, be limited given the good agreement between the enthalpies obtained and those expected for (Pro-Gly-Pro)$_9$. Lastly, it is probable that the end-blocks as part of the triblock copolymer have a higher T_m than that of free (Pro-Gly-Pro)$_9$ because of decreased entropy and because the end-blocks each only have one free terminus. Although the T_m may, thus, not be determined only by the length of the end-blocks, the variability in T_m is. Besides using the length of the end-blocks, the T_m could also be controlled by changing their amino acid composition. The relation between amino acid sequence and thermal stability of the triple helix has been described in detail[56].

The length of both the knot-forming and spacer regions in the triblock copolymers are exactly defined. In trimeric form the length of the knots is around 8 nm, as calculated from the known size of (Pro-Pro-Gly)$_{10}$ trimers[57]. On the basis of data from a set of 33 unfolded proteins[58], the radius of gyration for the present mid-blocks is expected to be around 7 nm, and the root-mean-square distance between the end-blocks on opposite sides of the molecule is around 18 nm (or slightly bigger owing to the tendency of moderately proline-rich sequences to form transient local polyproline-II-like structure elements). In a homogeneous gel, such as in the smooth interfaces observed with SEM, this size range seems suitable for slow release of, for example, protein drugs. The trial study on release of lysozyme indeed suggests these gels could be further developed for such applications. For example, longer or shorter mid-blocks could be used to alter the distance between the cross-links at both ends of the polymer and, thus, indirectly, the cross-link density and release characteristics of the gel.

The erosion rate of the gels may also need to be controlled. Coiled-coil-based triblock polymer gels in excess fluid have been shown to erode quickly at a constant rate from the surface, owing to loop formation between the end-blocks and the release of disengaged clusters from the network[38]. In principle, loops can also form in our system, depending on the length of the mid-block. It took two days for ~70 % of the 1 mm thick TR4T gel to erode at 20 °C. Although this seems fairly stable, applications requiring release over extended periods would still require chemical or enzymatic cross-linking. Shen et al[38]. showed it was possible to reduce the erosion rate of coiled-coil-based triblock gels by more than one hundredfold by employing dissimilar end-blocks that do not readily associate with each other and, thus, cannot form loops. A means of achieving this in our collagen-based system may possibly be found in the work of Gauba and Hartgerink, who demonstrated preferential self-assembly of heterotrimers over homotrimers upon mixing oppositely charged chemosynthetic collagenous peptides[59, 60]. To provide the desired charge, Pro or Hyp residues in each peptide were replaced with either Asp, Glu, Lys or Arg. Applying this approach to our system, triblocks with positively charged end-blocks could be mixed with triblocks that have negatively charged end-blocks. Loop formation would be prevented by mutual charge repulsion of the end-blocks, while a gel could still form through the formation of heterotrimers between oppositely charged end-blocks of different molecules. Moreover, different types of mid-blocks could be combined with each type of end-block, allowing a unique level of control over the molecular architecture of the gels. This could be of interest particularly for biomedical applications, as there is considerable freedom to insert or, conversely, avoid specific cell-binding, protein-binding, protease-sensitive, or bioactive amino acid sequences in the mid-blocks.

2.5 CONCLUSIONS

Triblock copolymers consisting of terminal $(\text{Pro-Gly-Pro})_9$ blocks and a random coil spacer have been efficiently produced as extracellular proteins in recombinant yeast. Specifically the end-blocks form collagen-like triple helices, contrasting traditional gelatins where cross-links form along the entire chain of each molecule. Unlike animal gelatins, the dynamic elasticity of the gels is independent of the thermal history, and the T_m is inherently defined by the length of the end-blocks. Together, these characteristics offer novel possibilities for controlled drug release and other biomedical applications.

ACKNOWLEDGEMENT

We thank Jeroen Bastiaans, Claudia Patrignani, Bas Simon, Patrick van Doeveren, Martijn Hendriksen, Adrie Westphal, Paulina Skrzeszewska, and Jacqueline Donkers (all from Wageningen UR) for technical assistance, Frans van Hoesel (Centre for High Performance Computing and Visualisation, University of Groningen) for generating the 3D impression of the gel structure, and Hans Peter Bächinger (Oregon Health and Science University, Department of Biochemistry & Molecular Biology) for kindly providing CD spectral data of Ac-(Gly-Pro-Pro)$_{10}$-NH$_2$.

This research was financially supported in part by the Netherlands Ministry of Economic Affairs and The B-Basic partner organisations (www.b-basic.nl) through B-Basic, a public-private NWO-ACTS programme (ACTS = Advanced Chemical Technologies for Sustainability).

REFERENCES AND NOTES

1. N. M. Sangeetha and U. Maitra, *Chem Soc Rev*, 2005, **34**, 821-836.
2. D. Hutmacher, M. B. Hurzeler and H. Schliephake, *Int J Oral Maxillofac Implants*, 1996, **11**, 667-678.
3. A. Asghar and R. L. Henrickson, *Adv Food Res*, 1982, **28**, 231-372.
4. J. Myllyharju and K. I. Kivirikko, *Ann Med*, 2001, **33**, 7-21.
5. J. Baez, D. Olsen and J. W. Polarek, *Appl Microbiol Biotechnol*, 2005, **69**, 245-252.
6. R. Rohanizadeh, M. V. Swain and R. S. Mason, *J Mater Sci Mater Med*, 2007, **19**, 1173-1182.
7. R. L. Watts, C. D. Raiser, N. P. Stover, M. L. Cornfeldt, A. W. Schweikert, R. C. Allen, T. Subramanian, D. Doudet, C. R. Honey and R. A. Bakay, *J Neural Transm Suppl*, 2003, 215-227.
8. F. R. Huss, J. P. Junker, H. Johnson and G. Kratz, *J Plast Reconstr Aesthet Surg*, 2007, **60**, 543-555.
9. E. Mattens, P. Engels, R. Hamerlijnck, J. Kelder, M. Schepens, J. de Valois, F. Vermeulen and L. Wijers, *Cardiovasc Surg*, 1999, **7**, 432-435.
10. S. Takebayashi, M. Hosaka, Y. Kubota, E. Ishizuka, A. Iwasaki and S. Matsubara, *J Urol*, 1998, **159**, 696-701.
11. D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen and J. Polarek, *Adv Drug Deliv Rev*, 2003, **55**, 1547-1567.
12. *US Pat.*, 20060241032, 2006.
13. T. Chen, R. Janjua, M. K. McDermott, S. L. Bernstein, S. M. Steidl and G. F. Payne, *J Biomed Mater Res B Appl Biomater*, 2006, **77**, 416-422.
14. *US Pat.*, 20050101531, 2005.
15. *US Pat.*, 20050119170, 2005.
16. M. Sutter, J. Siepmann, W. E. Hennink and W. Jiskoot, *J Control Release*, 2007, **119**, 301-312.
17. B. Balakrishnan, M. Mohanty, P. R. Umashankar and A. Jayakrishnan, *Biomaterials*, 2005, **26**, 6335-6342.
18. *US Pat.*, 20060204511, 2006.
19. S. Hassoun and A. Sabbah, *Allerg Immunol (Paris)*, 1998, **30**, 80-82.
20. D. Olsen, J. Jiang, R. Chang, R. Duffy, M. Sakaguchi, S. Leigh, R. Lundgard, J. Ju, F. Buschman, V. Truong-Le, B. Pham and J. W. Polarek, *Protein Expr Purif*, 2005, **40**, 346-357.
21. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
22. I. Goldberg, A. J. Salerno, T. Patterson and J. I. Williams, *Gene*, 1989, **80**, 305-314.
23. J. Capello, *Trends Biotechnol*, 1990, **8**, 309-311.
24. T. Kajino, H. Takahashi, M. Hirai and Y. Yamada, *Appl Environ Microbiol*, 2000, **66**, 304-309.
25. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.

26. J. M. Saddler and P. J. Horsey, *Anaesthesia*, 1987, **42**, 998-1004.
27. A. Vuorela, J. Myllyharju, R. Nissi, T. Pihlajaniemi and K. I. Kivirikko, *Embo J*, 1997, **16**, 6702-6712.
28. D. R. Olsen, S. D. Leigh, R. Chang, H. McMullin, W. Ong, E. Tai, G. Chisholm, D. E. Birk, R. A. Berg, R. A. Hitzeman and P. D. Toman, *J Biol Chem*, 2001, **276**, 24038-24043.
29. J. Myllyharju, *Matrix Biol*, 2003, **22**, 15-24.
30. P. D. Toman, G. Chisholm, H. McMullin, L. M. Giere, D. R. Olsen, R. J. Kovach, S. D. Leigh, B. E. Fong, R. Chang, G. A. Daniels, R. A. Berg and R. A. Hitzeman, *J Biol Chem*, 2000, **275**, 23303-23309.
31. J. Engel, J. Kurtz, E. Katchalski and A. Berger, *J Mol Biol*, 1966, **17**, 255-272.
32. J. Engel, H. T. Chen, D. J. Prockop and H. Klump, *Biopolymers*, 1977, **16**, 601-622.
33. S. Frank, R. A. Kammerer, D. Mechling, T. Schulthess, R. Landwehr, J. Bann, Y. Guo, A. Lustig, H. P. Bächinger and J. Engel, *J Mol Biol*, 2001, **308**, 1081-1089.
34. D. I. Rozkiewicz, Y. Kraan, M. W. Werten, F. A. de Wolf, V. Subramaniam, B. J. Ravoo and D. N. Reinhoudt, *Chemistry*, 2006, **12**, 6290-6297.
35. W. A. Petka, J. L. Harden, K. P. McGrath, D. Wirtz and D. A. Tirrell, *Science*, 1998, **281**, 389-392.
36. C. Xu, V. Breedveld and J. Kopecek, *Biomacromolecules*, 2005, **6**, 1739-1749.
37. L. Mi, S. Fischer, B. Chung, S. Sundelacruz and J. L. Harden, *Biomacromolecules*, 2006, **7**, 38-47.
38. W. Shen, K. Zhang, J. A. Kornfield and D. A. Tirrell, *Nat Mater*, 2006, **5**, 153-158.
39. I. R. Wheeldon, S. C. Barton and S. Banta, *Biomacromolecules*, 2007, **8**, 2990-2994.
40. Y. Cao and H. Li, *Chem Commun (Camb)*, 2008, 4144-4146.
41. E. R. Wright, R. A. McMillan, A. Cooper, R. P. Apkarian and V. P. Conticello, *Adv Funct Mater*, 2002, **12**, 149-154.
42. F. Tanaka and W. H. Stockmayer, *Macromolecules*, 1994, **27**, 3943-3954
43. S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen and L. R. Pease, *Gene*, 1989, **77**, 51-59.
44. S. P. Chambers, S. E. Prior, D. A. Barstow and N. P. Minton, *Gene*, 1988, **68**, 139-149.
45. J. M. Cregg, K. J. Barringer, A. Y. Hessler and K. R. Madden, *Mol Cell Biol*, 1985, **5**, 3376-3385.
46. W. Zhang, M. A. Bevins, B. A. Plantz, L. A. Smith and M. M. Meagher, *Biotechnol Bioeng*, 2000, **70**, 1-8.
47. Y. Katakura, W. Zhang, G. Zhuang, T. Omasa, M. Kishimoto, Y. Goto and K. Suga, *J Ferment Bioeng*, 1998, **86**, 482-487.
48. K. Sutoh and H. Noda, *Biopolymers*, 1974, **13**, 2391-2404.
49. F. A. De Wolf and R. C. A. Keller, *Prog Colloid Polym Sci*, 1996, **102**, 9-14.
50. K. Mizuno, T. Hayashi, D. H. Peyton and H. P. Bächinger, *J Biol Chem*, 2004, **279**, 38072-38078.
51. P. Bruckner, B. Rutschmann, J. Engel and M. Rothe, *Helv Chim Acta*, 1975, **58**, 1276-1287.
52. A. A. Makarov, I. A. Adzhubei, Protasevich, II, V. M. Lobachov and N. G. Esipova, *J Protein Chem*, 1993, **12**, 85-91.

53. M. Djabourov, in *Progress in Biotechnology*, eds. W. I. Aalbersberg, R. J. Hamer, P. Jasperse, H. H. J. de Jongh, C. G. de Kruif, P. Walstra and F. A. de Wolf, Elsevier Science, Amsterdam, 2003, vol. 23, pp. 133-218.
54. R. A. Berg, B. R. Olsen and D. J. Prockop, *J Biol Chem*, 1970, **245**, 5759-5763.
55. K. Te Nijenhuis, *Colloid & Polymer Science*, 1981, **259**, 1017-1026.
56. A. V. Persikov, J. A. Ramshaw and B. Brodsky, *J Biol Chem*, 2005, **280**, 19343-19349.
57. R. Z. Kramer, L. Vitagliano, J. Bella, R. Berisio, L. Mazzarella, B. Brodsky, A. Zagari and H. M. Berman, *J Mol Biol*, 1998, **280**, 623-638.
58. N. C. Fitzkee and G. D. Rose, *Proc Natl Acad Sci U S A*, 2004, **101**, 12497-12502.
59. V. Gauba and J. D. Hartgerink, *J Am Chem Soc*, 2007, **129**, 2683-2690.
60. V. Gauba and J. D. Hartgerink, *J Am Chem Soc*, 2007, **129**, 15034-15041.

APENDIX 2.1

Table 2.2.1 Oligonucleotide sequences

Name	DNA sequence (5’-3’)
T-FW	GAGTCTCACCCGGTGCTC
T-RV	CCACCGGCTGGTCCGGGAGGACCCGGTGGTCCAGGTGGACCTGGTGGTCCTGGTGGACCAGGTGGACCAGGTGGACCAGGCGGACCGGGAGCACCGGGTGAGACTC
RA-FW	GCGCTCGAGAAAAGAGAGGCTGAAGCTGGTCCACCCGGTGAGTCACCAGGTCCTCAGCCTGGTGGTCCACAAAACCCAGGTTCCGGTGAAGGTCAGGGAAACGGTAACCCT
RA-RV	TGGGCCAGGGTTGGGGTTTGGTGGTTCACCGGAACCTGGAGAACCACCTCCTTGTGGTTGACCACCTTGAGATGGTCCATTCTTGTTAGGGTTACCGTTTCCCTGACCT
RB-FW	CAAACCCCAACCCTGGCCCACAGAACGGTCAAAAGCCTGGTGGTCAACAAAACGGTCCTGGTAATGGTCAACAAGAGGGAAACGGTCAACAAAACGGTGGT
RB-RV	GCGTCTGCAGTACGAATTCTATTAGCCACCGGCTGGCTGACCTGGAGGTTTTCCTGGTGGAGATCCAGGTCCGGACTGAGAACCACCACCGTTTTGTTGACCGTTTC

Chapter 3

Influence of chain length on gel-forming properties of telechelic collagen-inspired polymers

This chapter as been published in modifed form as: H. Teles, P. J. Skrzeszewska, M. W. T. Werten, J. van der Gucht, G. Eggink, and F. A. de Wolf, *Soft Matter*, 2010, DOI:10.1039/C0SM00175A.

SUMMARY

We studied the influence of polymer length on the formation of transient networks by telechelic polypeptides with ~ 2.3 kDa collagen-like triple helix-forming end-blocks and much longer random coil mid-blocks. We compared triblock copolymers with mid-blocks of ~400 and ~800 amino acids (~37 and ~74 kDa, respectively), both of which were secreted to high concentration by recombinant yeast cells. The longer triblock copolymers were produced and characterised for the first time at present. At the same molar concentration of protein and crosslink-forming end-blocks, the storage modulus of the longer polymers was much higher than of the shorter polymers. These results indicate that the elastic properties of the network are not only a function of concentration and temperature but also of polymer length. The experimentally obtained, temperature-dependent plateau storage modulus of all polymers were well described by an analytical model that was based on classical gel theory and accounted for the particular molecular structure of the gels, and the presence of loops and dangling ends. The model showed that the temperature at which 50 % of the end-blocks were involved in junctions was lower for the shorter polymers than for the longer polymers.

3.1 INTRODUCTION

Protein-based hydrogels have proven to be effective biomaterials in a variety of medical and pharmaceutical applications[1-3]. Despite the success of these materials, which are mostly derived from animal fibrillar proteins, their use is often determined by the characteristics of the available proteins, and it may be influenced by batch to batch (or source to source) variation, and the possible transmission of infectious agents associated with materials isolated from mammalian sources[4].

Traditional gelatin, commonly used in protein gels, consists of a multitude of different, partly degraded, and chemically modified fragments of animal collagen, resulting in gels with ill-defined molecular composition and complex network-forming properties. This complexity, in combination with the impossibility to change the molecular structure at will, limits the possibilities to determine structure-function relationships, and has prompted the need for exploration of synthetic strategies that allow exquisite control over the monomer sequence and polymer length[4].

In the past decade recombinant DNA techniques became a tool to develop genetically engineered protein polymers with defined composition and structure that offer safe biocompatibility and full control over the length and sequence of biopolymers. In chapter 2, we reported on the recombinant production and characterisation of a new class of gel-forming telechelic triblock gelatins with controllable and predictable properties. Gel formation is obtained by combining collagen-inspired (Pro-Gly-Pro)$_n$ end-blocks (T), inspired by and behaving like natural collagen, with highly hydrophilic random coil blocks defining the distance between the trimer forming T end-blocks. Contrary to natural gelatin gels, this new class of synthetic collagenous gels has a well-defined junction multiplicity of exactly three and predictable physical chemical properties. So far, two of such triblock copolymers, denoted as 'TP4T' and 'TR4T' have been produced and characterised (chapter 2 and reference[5]). These polymers consisted of 9 repeating units of (Pro-Gly-Pro) T blocks and a mid-block (P4 or R4) made of a tandem repeat of four highly hydrophilic 9 kDa blocks assuming random conformation: the polar ‘P’ block with a collagen-inspired designer sequence consisting of (Gly-Xaa-Yaa) repeats and the ‘R’ block with the same amino acid composition, but quasi random amino acid sequence. The mass of the entire TP4T or TR4T was 42 kDa. By changing the underlying DNA template, we doubled the length of the mid-block and here, we report for the first time on the biosynthesis and characterisation of the resulting ~78 kDa triblock copolymers, TP8T and TR8T. In our previous work, focussed on TR4T hydrogels[5], we showed that the network-forming properties are strongly dependent on concentration and temperature. By taking in consideration the formation of loops and dangling ends, we were able to develop an analytical model[5] that accurately described the network structure and truly reproduced the observed relationships between the storage modulus, temperature, concentration and time without a need for adjustable parameters other than quantities that could be verified by experimental measurements. In this chapter, besides concentration and temperature, we study the influence of mid-block length and amino acid sequence on the viscoelastic properties of the networks. The amino acid sequence appears to have an influence on the persistence length of the mid-block and, consequently, on the critical overlap concentration. We use our previous model to interpret the experimental data and show that small changes in polymer structure influence significantly the mechanical properties of the networks.

3.2 MATERIAL AND METHODS

Expression vectors and *Pichia pastoris* transformation

The construction of the vectors encoding the genes TP4T and TR4T are described in detail in chapter 2. The previously described vectors pMTL23-P4T and pMTL23-R4T, pMTL23-P4[6] and pMTL23-R4, and pCR4-TOPO-T were used to construct the genes encoding TP8T and TR8T. Vectors pMTL23-P4T and pMTL23-R4T were digested with *Dra*III (5´ to the P4T and R4T genes) and desphosphorylated. P4- and R4-DNA, obtained from *Dra*III/*Van*91I-digested pMTL23-P4 and pMTL23-R4, was inserted into the previously digested pMTL23-P4T and pMTL23-R4T, respectively, so as to yield pMTL23-P8T and pMTL23-R8T. These vectors were then digested with *Dra*III (5´ to the P8T and R8T genes) and desphosphorylated. A T block DNA, obtained from *Dra*III/*Van*91I-digested pCR4-TOPO-T, was inserted into the previously digested pMTL23-P8T and pMTL23-R8T, so as to yield vectors pMTL23-TP8T and pMTL23-TR8T. TP8T and TR8T fragments were then cloned into *P. pastoris* expression vector pPIC9 (Invitrogen) using XhoI/EcoRI. The entire P4[6], R4 (chapter 2), TP4T (chapter 2), TR4T (chapter 2), TP8T and TR8T DNA sequences (and translated amino acid sequences) have been deposited in the GenBank under acession numbers ACF33476-ACF33481. The expression vectors were linearised with *Sal*I and transformed into the *his4* GS115[7] strain. Transformation and strain selection of transformants was as described earlier[8].

Fermentation of *P. pastoris*

Fermentation of *P. pastoris* strains was performed as described by Zhang et al[9] . Strains were cultured in minimum basal salts medium[9] in a 2.5-liter Bioflo 3000 fermenter (New Brunswick Scientific). The fermentation was run at 30 °C and the pH was kept at 3 by addition of ammonium hydroxide. The methanol level in the fermentation broth was kept constant at ~0.2 % (w/v) by using a homemade semiconductor gas sensor-controller, similar to that described by Katakura et al[10]. The methanol fed-batch phase was prolonged for approximately 2-3 days. The cells were removed by centrifugation for 30 min at 10,000x g in a SLA-3000 rotor (Sorval), followed by microfiltration of the supernatant.

Protein Purification

Protein purification was done by ammonium sulphate precipitation as described previously in chapter 2.

SDS-page

Fermentation supernatants and pure products were analysed by SDS-page using the NuPAGE Novex system (Invitrogen), with 10 % Bis-Tris gels, MES SDS running buffer and SeeBlue Plus2 pre-stained molecular mass markers. Gels were stained in

Coomassie SimplyBlue SafeStain (Invitrogen) and distained in MQ water. Densitometric quantification of pure TP8T or TR8T standards, ranging from 2 µg to 25 µg micrograms, and culture supernatants was done with Bio-Rad GS-800 densitometer and analised using Quantity One computer software. N-terminal sequencing analysis was determined by Edman degradation as described previously[8]. Protein sequencing was preformed by Midwest Ananlytical (St. Louis, Mo).

Mass spectroscopy

MALDI-TOF was performed using an Ultraflex mass spectrometer (Bruker). Samples were prepared by the dried droplet method on a 600 µm AnchorChip target (Bruker), using 5 mg/ml 2,5-dihydroxyacetophenone, 1.5 mg/ml diammonium hydrogen citrate, 25 % (v/v) ethanol and 1 % (v/v) trifluoroacetic acid as matrix. Measurements (20 Hz) were made in the positive, linear mode, with the following parameters: ion source 1, 20000 V; ion source 2, 18450 V; lens, 5000 V; pulsed ion extraction, 550 ns. Protein Calibration Standard II (Bruker) was used for external calibration.

Rheology

The rheological measurements where preformed with an Anton Paar Physica MCR301 rheometer, equipped with a cone and plate geometry of 50 mm diameter. Temperature control was insured by a Peltier system, which allowed fast heating and cooling. Protein solutions of different molar concentrations were prepared in phosphate buffer (10 mM, pH 7), heated for half an hour at 50 °C and then introduced in the pre-heated geometry. The system was subsequently quenched to 20 °C and gel formation was followed for 15 hours. A solvent trap was used to prevent evaporation. Gelation was monitored by measuring the storage modulus (G'), as well as the loss modulus (G'') at a frequency of 1 Hz and a controlled strain of 1 %. Two types of rheological measurements were carried out at different temperatures between 20 and 40 °C). At each temperature the system was equilibrated for 5h before doing the measurements. First, viscoelastic characterization of the equilibrated gel was performed in the frequency range 0.001-20 Hz (0.00628-125 rad/s) and a deformation amplitude of 1 %. Second, to explore the behavior of the system below 0.01 rad/s, creep experiments were carried out and converted to the frequency domain using methods described by Ferry[11]. For TP4T, the applied stress was varied between 5 Pa and 20 Pa and for TP8T and TR8T, between 5 Pa and 300 Pa. The stress values were chosen not to go beyond the linear regime. The deformation and recovery phases were followed each for a period of 1800 s. The TR4T experimental data presented in this chapter have been acquired previously[5].

Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry experiments were performed with a MicroCal VP-DSC instrument. 0.51 ml degassed protein solutions (2.4 mM) prepared in phosphate buffer 10 mM, pH 7) were loaded at 50 °C into the calorimeter. For each experiment, the protein solutions were equilibrated for 10 h at 20 °C to allow complete helix formation. Then, the temperature was raised from 20 to 65 °C at a scan rate of 15 °C/h. At the protein concentration used both melting temperature and enthalpy are scan rate independent[12]. After recording the DSC trace, the sample was cooled to 20 °C and left for 10 h for helix re-anneaniling. The DSC trace was then recorded again in an identical manner. The calorimetric transition enthalpy was obtained by integration of the area under the excess heat capacity peak. The flooring of the thermogram was done by fitting the thermogram baseline with a polynomial function of degree three and the area under the endotherm was determined using the Simpson's rule.

Dynamic Light Scatering

Protein solutions of 0.5 mg/ml prepared in phosphate buffer (10 mM, pH 7) were pipetted into light scattering cuvettes of 45 µl (3 mm×3 mm×5 mm) The data were acquired at 20 °C using a Malvern Zetasizer Instrument (Nano series). The light source was a 4 mW He-Ne laser with wavelength 633 nm. The scattering angle θ was fixed at 173°.

3.3 EXPERIMENTAL RESULTS AND DISCUSSION

Biosynthesis and molecular characterisation of the protein polymers

P. pastoris strains containing the genes encoding the ~78 kDa collagen-like triblock-copolymers targeted to the extra-cellular medium, TP8T and TR8T, were cultured in bioreactors and the proteins were purified from the fermentation supernatant. Cell-free TP8T and TR8T fermentation broths and purified products were analysed by SDS-PAGE (Figure 3.1a). All purified proteins migrated as single bands indicating a high purity and intactness, although TP8T and TR8T showed anomalous migration behaviour in SDS-PAGE, similar to the migration behaviour found previously for P4 (chapter 2 and reference[6]), TP4T and TR4T (chapter 2). The molecular weights of TP8T and TR8T were determined by MALDI-TOF MS. Figure 3.1b shows that the experimental values are in good agreement with the expected value of 78176 DA for both TP8T and TR8T. Furthermore, N-terminal sequencing of the SDS-PAGE bands confirmed the identity of the proteins.

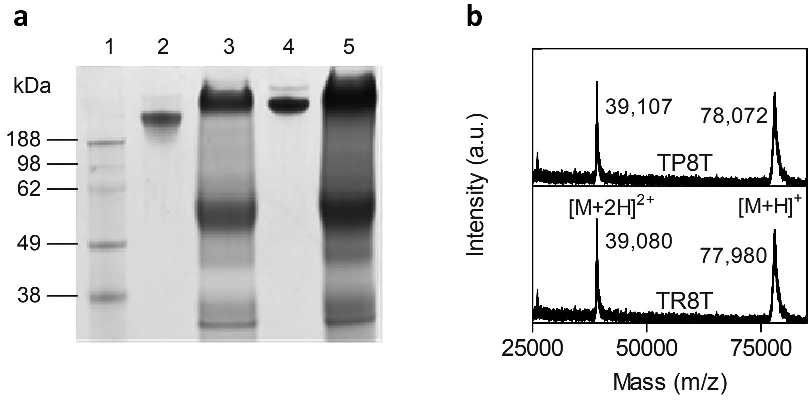

Figure 3.1 Production of protein polymers (a) SDS-PAGE: lane 1, molecular weight marker; lane 2, purified TR8T; lane 3, cell-free TR8T fermentation broth; lane 4, purified TP8T; lane 5, cell-free TP8T fermentation broth. 15 µl of cell-free fermentation broth and ~20 µg of purified protein was applied; (b) MALDI-TOF of purified TP8T and TR8T. Single and doubly charged molecular ions are indicated.

The productivity of TP8T and TR8T was estimated to be 1-3 g/l of clarified broth. These values are very similar to the production yields previously reported for P4[6], TP4T and TR4T (chapter 2). We can deduce from these results that an increase in total protein size did not have an influence on the proteins' intactness and yield.

Enthalpy and melting temperature

Endothermic transitions occurred upon melting TP8T and TR8T with a ΔCp maximum at 315.6 K and 315.3 K, respectively (Figure 3.2a and 3.2b). The calorimetric transition enthalpies were calculated by integration of the area under the excess heat capacity peaks. The calculated enthalpies were 229 ± 20 and 202 ± 11 kJ/mol triple helix for TP8T and TR8T, respectively. The experimental enthalpy values measured correlate well with the expected melting enthalpy value of 223 kJ/mol for a $(\text{Pro-Gly-Pro})_9$ end block calculated on the basis of the results obtained by Frank et al.[13] and Engel et al.[14]. These results are consistent with the earlier reported values of T_m and enthalpy for TP4T and TR4T (chapter 2). This confirms our previous findings that the trimer-forming end-blocks (T) are solely responsible for trimerization.

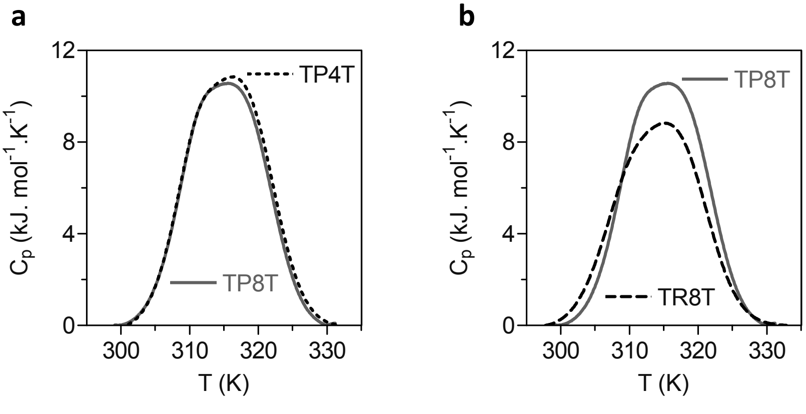

Figure 3.2 (a) Thermogram of TP8T and TP4T and (b) Thermogram of TP8T and TR8T. Protein concentration 2.4 mM.

Hydrogel formation

In line with our previous studies (chapter 2 and reference[5]), the gel-forming behaviour of TP4T, TR4T[5], TP8T and TR8T was investigated by means of dynamic rheology. Several concentrations of the four different protein gels were investigated. At 20 °C, a physical gel is formed due to triple helix formation exclusively by the collagen-like end-blocks (T). In the beginning of the gelation process viscous properties are predominant but with time, as the protein network develops, the elastic properties prevail. A steady-state storage modulus is reached after 5-6 hours. All four different proteins show this same course of network formation. Figure 3.3a shows the development of the storage (G’) and loss (G’’) modulus for a 1.2 mM solution of TP8T and TR8T. Similar data can be found for TP4T and TR4T in chapter 2 and reference[5].

After the steady state value of G’ was reached in TP4T, TP8T and TR8T gels, the G’ and G’’ values were measured as a function of frequency. Within the available frequency range, the elastic behaviour (G’) dominates over the viscous behaviour (G’’). The storage modulus does not depend on the frequency in this range, while the loss modulus decreases with frequency and reaches a minimum around 0.1 rad/s and then increases again (Figure 3.3b). The same behaviour has been reported before for TR4T[5] gels.

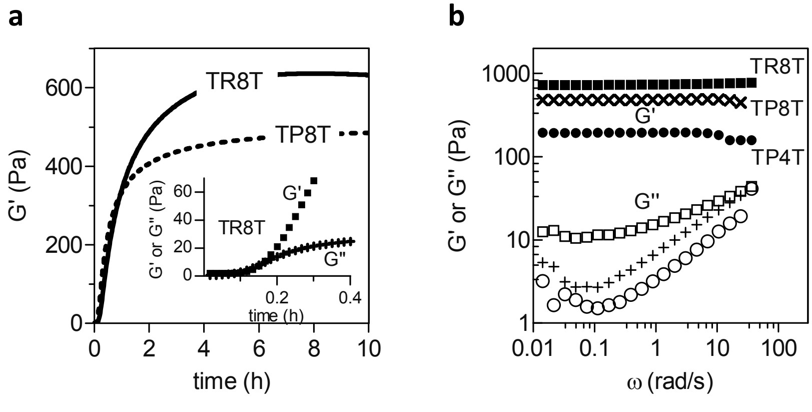

Figure 3.3 (a) Storage modulus for 1.2 mM TR8T (—) and TP8T (- -) as a function of age (1Hz, γ=1 %). Inset shows a zoom of the initial stages of TR8T gelation: G’ (■), G’’ (+). (b) Frequency sweeps at 293K for 1.2 mM TR8T and TP8T and 2.3 mM for TP4T.

To obtain insight into the gel system at frequencies below 0.01 rad/s, creep measurements were preformed. A fixed stress was applied to the gel after it reached a steady G’ value and the resulting deformation was followed as a function of time. We converted the creep results to the frequency domain[5] and combined them with the oscillatory shear data. Figure 3.4 shows creeps curves and the resulting frequency-dependent storage and loss modulus for several concentrations of TP4T, TP8T and TR8T. Similar results were found for TR4T[5]. Both frequency-dependent and creep experiments allowed us to determine the plateau storage modulus G_0 and relaxation time τ at different concentrations (symbols in Figure 3.5a-b respectively). The plateau storage modulus of all proteins studied depends strongly on concentration[5]. Furthermore, the longer versions of the collagen-like protein polymers, i.e., TP8T and TR8T have a considerably higher plateau storage modulus (G_0) than their shorter counterparts, TP4T and TR4T (Figure 3.5a). While a storage modulus of ~ 700 Pa is obtained at a protein concentration of 1.2 mM with TR8T, a comparable value is only reached at a concentration of ~2 mM with TR4T gels. This difference is most probably related to the fact that longer chains are less likely to form intramolecular loops, and therefore lead to a higher density of trimolecular (network-building) junctions at the same polymer concentration[5]. In an attempt to provide a consistent and more thorough interpretation of these observations, we use a model that we have developed[5] for networks of telechelic polymers with trimer- forming end-blocks.

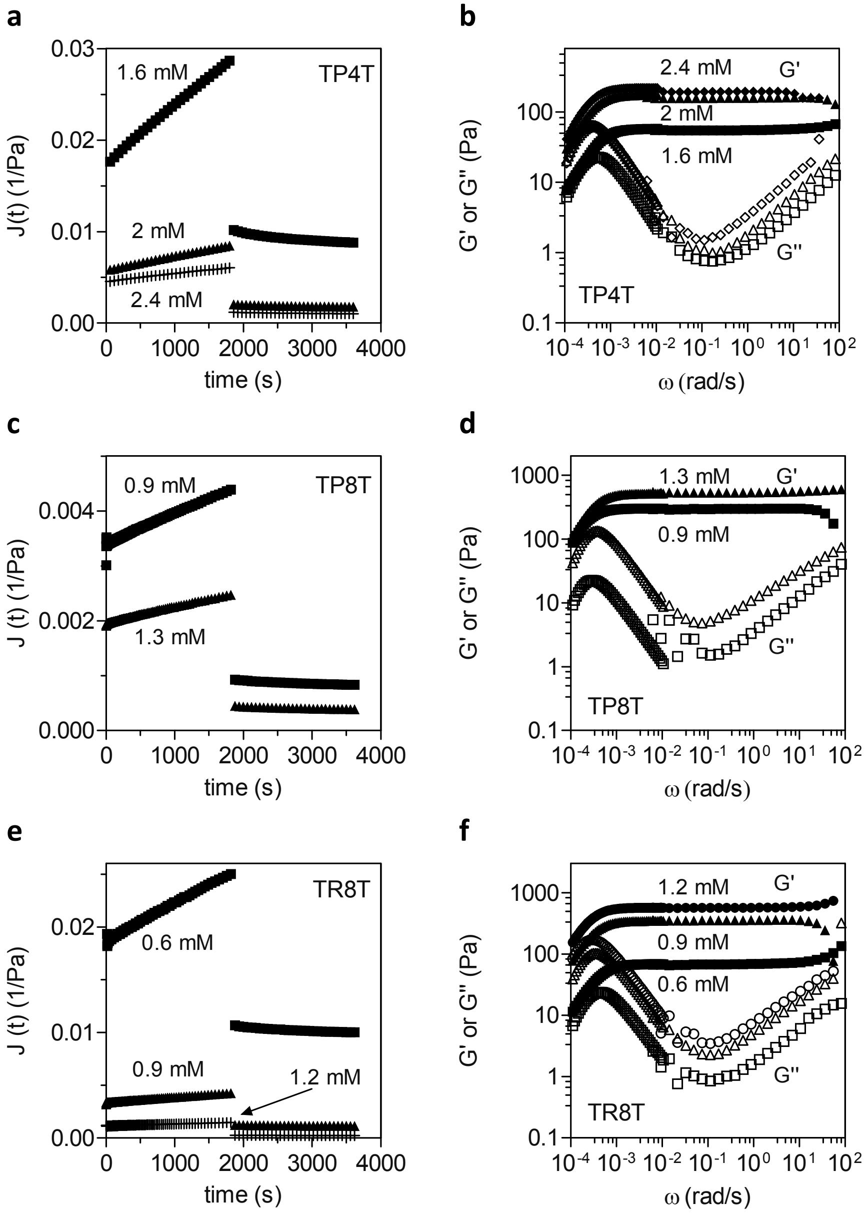

Figure 3.4 Creep experiments at 293 K for (a) TP4T, (c) TP8T and (e) TR8T at different concentrations as indicated; Combined frequency sweeps and converted creep experiments at 293K for (b) TP4T, (d) TP8T and (f) TR8T at different concentrations as indicated.

Modelling of the network

The model was developed based on classical Flory[15]-Stockmayer[16] theory of gel formation and allows calculation of the complete (linear) viscoelastic behaviour of the network. The model takes into account the equilibrium between helix formation and free ends by the trimer forming end-blocks (T)[5, 14]. The formation of helices can occur in two ways: a triple helix formed by three T end-blocks coming from three different chains (junction) or, by two of the three T end-blocks coming from the same chain (loop) and one from a different chain. Only junction points with all three branches linked to the gel network contribute effectively to the elastic properties, while loops and dangling ends act as gel stoppers. The junction/loop ratio increases with protein concentration and molecular size[5]. The concentration of chains involved in loops (C_L) can be expressed as a function of concentration of junctions (C_J), overlap concentration (C^*) and concentration of total protein (C_T)[5]:

$$C_L = \frac{C_J}{2C_T V_{coil}} \approx \frac{C_J C^*}{2C_T} \tag{3.1}$$

where C^*, is related to the dimension of the chains and can be estimated as $C^* \sim 1/(4/3\pi R_g^3)$, where R_g is the radius of gyration of a polymer coil of a certain volume V.

The equation to calculate the elastic modulus (G_0) was derived from the classical gelation theory[15, 16]:

$$G_0 = \frac{F\left[C_T(2p-1)^3 RT\right]}{p^2} \tag{3.2}$$

where RT is the thermal energy per mole, F is the front factor, taken to be 0.5[5], and $p=3C_J/2C_T$ is the probability of a T end group to be involved in a junction point.

Stress relaxation occurs when an elastically active chain breaks. Here, this happens when a triple helix (junction point) dissociates[5]. The relaxation time is given by the following equation:

$$\tau = \tau_0 (2p - 1) \qquad . \qquad (3.3)$$

where τ_0 is the life time of a single helix. Dissociation of helices is also a function of temperature. The higher the temperature the faster the helices dissociate. This temperature dependence can be expressed by the Van ’t Hoff’s equation:

$$K_H = K_{H,0} \exp\left[\frac{-\Delta H}{RT}\right] \qquad (3.4)$$

where ΔH is the molar enthalpy of helix formation. Since helix formation is a exothermic process (ΔH <0), $K_{H,}$ the triple helix association equilibrium constant, decreases with increasing temperature[5]. The shift in K_H brings changes in the concentration of junctions (C_J), loops ($C_{L)}$ and free ends (C_F) and thus changes in the storage modulus, the relation time and the viscosity[5].

For the model calculations with mid-block other than R4 it was necessary to recalculate the overlap concentration (C*) which is related to the dimensions of the protein chain. All the other parameters, determined experimentally, were kept constant with respect to our former work. For more details on the model calculations please see reference[5].

To have a good estimation of the dimension of the protein chains in solution we used dynamic light scattering to measure the size of P4 and R4 blocks. The results obtained indicate an ~8 % lower hydrodynamic radius (R_h) for the P4 block (5.7 ± 0.3 nm) as compared to the R4 block (6.2 ± 0.4 nm). The measured hydrodynamic radius of R4 is smaller than the radius of gyration (R_g) of ~7 nm estimated earlier[5] based on a master curve for R_g as a function of the number of residues proposed by Fitzkee et al.[17].

However, the hydrodynamic radius is usually slightly lower than the radius of gyration for random coil proteins[18, 19]. Assuming R_g to be ~11 % higher than R_h, the calculated overlap concentrations for TP4T and TR4T were ~ 1.6 mM and ~1.1 mM[5], respectively.
We assumed the R_g to scale with the residue number as $R_g \sim N^{0.6}$ for polymers in good solvent[17, 19]. Accordingly, we estimated the R_g of the polypeptides with the longer mid-blocks, P8 and R8, to be ~10 and ~11 nm, respectively, thus maintaining the same size ratio between the P8 and R8 blocks as between the P4 and R4 blocks. The corresponding values of the overlap concentrations (C^*) of TP8T and TR8T were ~0.38 mM and ~0.3 mM, respectively. Having determined the overlap concentration (C^*) of all four proteins, we could model the network. Figures 3.5a-e show a very good agreement between the experimental data (symbols) and model calculations (solid lines) for all proteins.

It is apparent from the model that the polymers with longer mid-blocks have a higher storage modulus because a longer mid-block leads to a bigger radius of gyration (R_g) and thus to a lower overlap concentration. A bigger R_g decreases the probability that two end-blocks from the same molecule associate with each other, i.e., form a loop. The consequence of fewer loops in the system is a higher storage modulus[5]. Figure 3.6a shows the dependency of the fraction of junctions (p_J) and loops (p_L) on the concentration for all four proteins, based on model calculation. In this figure it is clearly shown that the fraction of loops present in the system is considerably higher for the shorter triblock copolymers, TP4T and TR4T at each concentration.

In addition, we observed that the R series, TR4T and TR8T, also show a higher storage modulus than their corresponding P series counterparts, TP4T and TP8T (Figure 3.5a). The difference is more pronounced for the shorter versions, TP4T and TR4T. Although, P and R blocks are both random coils with exactly the same amino acid composition, their amino acid sequence is different. Fitzkee et al[17] have shown that even a polypeptide chain that assume a random coil conformation still has locally folded conformations that contribute to the overall flexibility of the chain. Apparently, this leads to a smaller radius of gyration for the P mid-blocks than for the R mid-blocks. This effect is stronger for the proteins with the shorter mid-blocks, because the fraction of loops decreases strongly with increasing polymer length (Figure 3.6a).

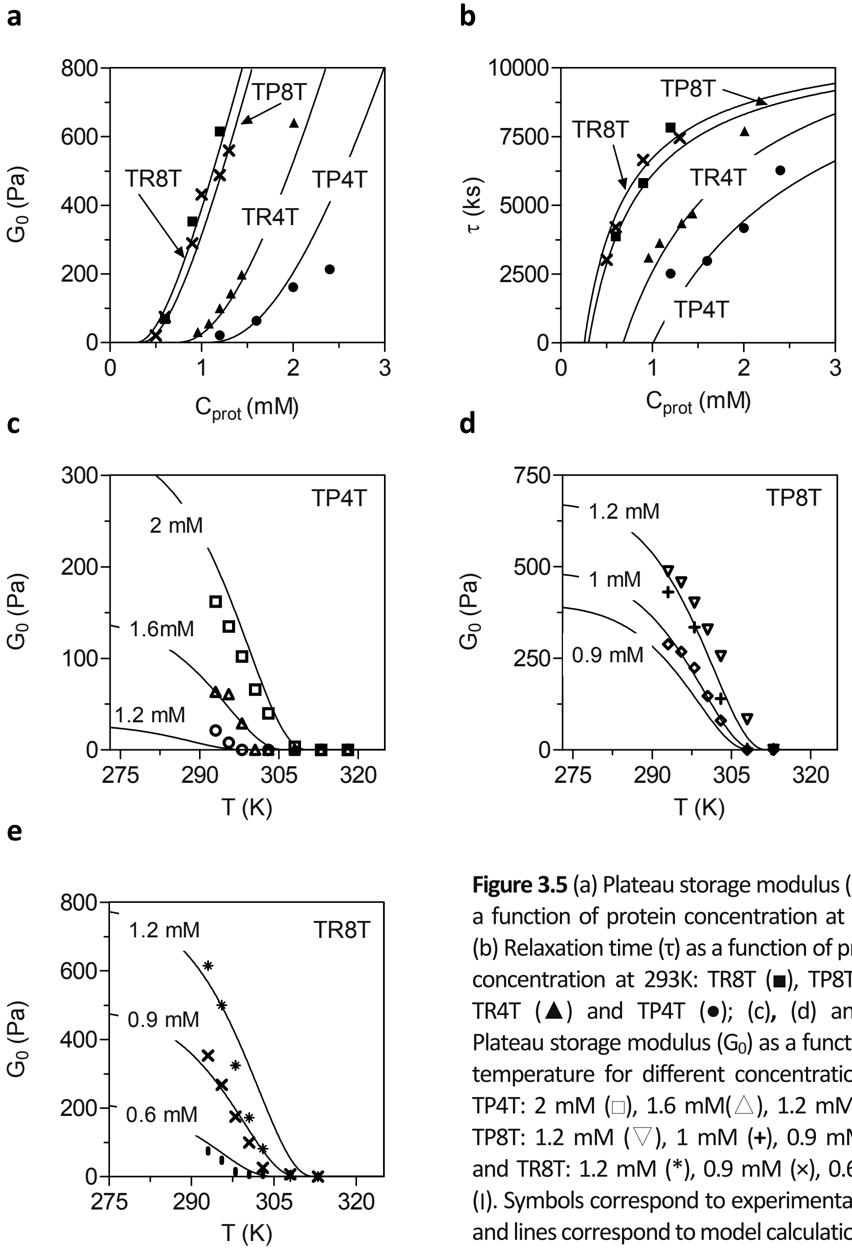

Figure 3.5 (a) Plateau storage modulus (G_0) as a function of protein concentration at 293K; (b) Relaxation time (τ) as a function of protein concentration at 293K: TR8T (■), TP8T (**X**) , TR4T (▲) and TP4T (●); (c)**,** (d) and (e) Plateau storage modulus (G_0) as a function of temperature for different concentrations of TP4T: 2 mM (□), 1.6 mM(△), 1.2 mM (○) ; TP8T: 1.2 mM (▽), 1 mM (+), 0.9 mM (◊), and TR8T: 1.2 mM (*), 0.9 mM (×), 0.6 mM (I). Symbols correspond to experimental data and lines correspond to model calculations.

Melting behaviour of the network

In order to compare the melting behaviour of the four proteins, creep and frequency sweeps were carried out also at higher temperatures (293-323 K). Figures 3.5c-e show the plateau storage modulus (G_0) of TP4T, TP8T, ad TR8T as a function of temperature, at different concentrations. Contrary to what emerged from DSC measurements (Figure 3.2) the temperature at which the G_0 value approaches zero and the gel completely loses its elastic properties, varies with the length of the mid-block. For the same molar concentration, e.g. 1.2 mM, this temperature is ~298 K for TP4T and ~313 K for TP8T. Those temperatures should correspond approximately to the temperature at which 50 % of all the T end groups available in the system (i.e. twice the amount of gelatin molecules) are involved in trimolecular junctions and thus the junction probability $p = 3C_J/2C_T = 0.5$.

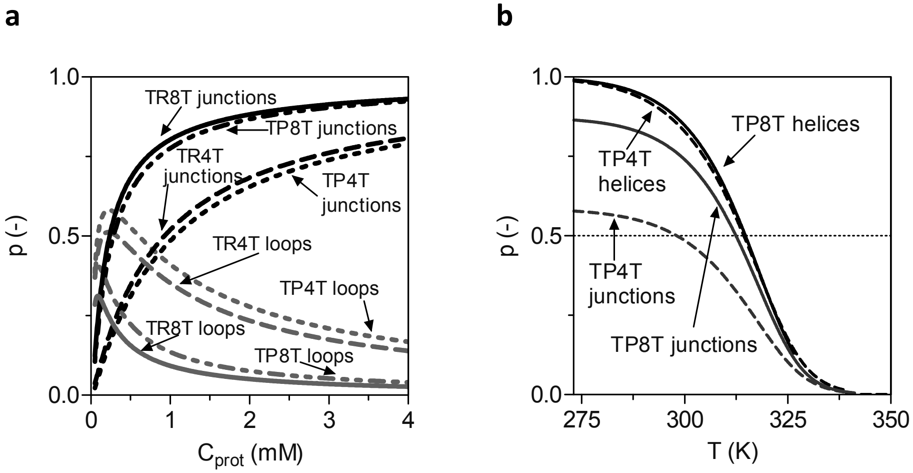

Figure 3.6 (a) Dependency of fraction of junctions (p_J) and loops (p_L) on protein concentration; (b) Dependency of fraction of triple helices (p_H) and junctions (p_J) on temperature for TP4T and TP8T at a protein concentration of 1.2 mM.

In Figure 3.6b we plot the fraction of triple helices (p_H) and fraction of junctions (p_J) obtained from model calculations as a function of temperature for TP4T and TP8T at a concentration of 1.2 mM. The total amount of triple helices (junctions and loops together) does not depend on the length of the mid-block and reaches ~50 % at ~315 K for both TP4T and TP8T. However, the fraction of junctions is significantly higher for molecules with a longer mid-block. The 50 % value is reached at 297 K for TP4T and at 313 K for TP8T in good agreement with the temperature dependence of G_0 shown in Figure 3.5c-d, and discussed above.

3.4 CONCLUDING REMARKS

In order to explore the relationship between the mid-block size and hydrogel-forming properties of telechelic polypeptides with trimer forming collagen-like end-blocks four versions of this class of polymers, differing only in their mid-block length or amino acid sequence, were studied. The temperature-dependent measurement of the storage modulus (reflecting the elasticity of the gels) showed that, besides polymer concentration[5], the length of the random coil mid-blocks is also a predominant factor governing the viscoelastic properties of the hydrogels. Due to loop formation by the trimer-forming end-blocks, the gel strength depends not only on the helix content of the system but also on the ratio between loops and junctions. Longer polymer chains render a higher gel strength, because the probability is lower that their two end-blocks associate with each other to form a loop. A consequence of fewer loops in the system is a higher number of active chains effectively contributing to the elastic properties of the network. The results were supported by an analytical model based on classical gel theory. The model calculations were in good agreement with the experimental data showing, once more, that the well-defined multiplicity of the network of this particular class of hydrogels allows the prediction of the complete viscoelastic behaviour of the network, when the polymer structure is well known.

Our results suggest that, by controlling the structure of the present type of hydrogel-forming polymers through genetic engineering their physical-chemical properties can not only be controlled and predicted but also developed in order to match a variety of different applications.

ACKNOWLEDGEMENT

This research was financially supported in part by The Netherlands Ministry of Economic Affairs and the B-Basic partner organisations (www.b-basic.nl) through B-basic, a public-private NWO-ACTS programme (ACTS=Advanced Chemical Technologies for Sustainability).

REFERENCES

1. C. Yang, P. J. Hillas, J. A. Baez, M. Nokelainen, J. Balan, J. Tang, R. Spiro and J. W. Polarek, *BioDrugs*, 2004, **18**, 103-119.
2. D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen and J. Polarek, *Adv Drug Deliv Rev*, 2003, **55**, 1547-1567.
3. C. H. Lee, A. Singla and Y. Lee, *Int J Pharm*, 2001, **221**, 1-22.
4. R. Langer and D. A. Tirrell, *Nature*, 2004, **428**, 487-492.
5. P. J. Skrzeszewska, F. A. de Wolf, M. W. T. Werten, A. P. H. A. Moers, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2009, **5**, 2057-2062.
6. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
7. J. M. Cregg, K. J. Barringer, A. Y. Hessler and K. R. Madden, *Mol Cell Biol*, 1985, **5**, 3376-3385.
8. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
9. W. Zhang, M. A. Bevins, B. A. Plantz, L. A. Smith and M. M. Meagher, *Biotechnol Bioeng*, 2000, **70**, 1-8.
10. Y. Katakura, W. Zhang, G. Zhuang, T. Omasa, M. Kishimoto, Y. Goto and K. Suga, *J. Ferment. Bioeng.*, 1998, **86**, 482-487.
11. J. D. Ferry, *Viscoelastic Properties of Polymers*, Wiley, New York, 1980.
12. P. J. Skrzeszewska, F. A. De Wolf, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2010, **6**, 416-422.
13. S. Frank, R. A. Kammerer, D. Mechling, T. Schulthess, R. Landwehr, J. Bann, Y. Guo, A. Lustig, H. P. Bachinger and J. Engel, *J Mol Biol*, 2001, **308**, 1081-1089.
14. J. Engel, H.-T. Chen, D. J. Prockop and H. Klump, *Biopolymers* 1977, **16**, 601-622.
15. P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press, New York, 1953.
16. W. H. Stockmayer, *J. Chem. Phys.*, 1944, **12**, 125-136.
17. N. C. Fitzkee and G. D. Rose, *Proc Natl Acad Sci U S A*, 2004, **101**, 12497-12502.
18. D. K. Wilkins, S. B. Grimshaw, V. Receveur, C. M. Dobson, J. A. Jones and L. J. Smith, *Biochemistry*, 1999, **38**, 16424-16431.
19. C. Tanford, *Physical Chemistry of Macromolecules*, John Wiley & Sons, Inc., New York, 1961.

Chapter 4

Hydrogels of collagen-inspired telechelic triblock copolymers for the sustained release of proteins

This chapter as been accepted for publication in modified form in the *Journal of Controlled Release* as: H. Teles, T. Vermonden, G. Eggink, W.E. Hennink, F.A. de Wolf, 2010.

SUMMARY

In this chapter we study the erosion and protein release behaviour of two hydrogel-forming collagen-inspired triblock copolymers, differing only in their mid-block length (molecular weights ~37 kDa and ~73 kDa). These polymers, produced as heterologous protein in recombinant yeast, are made of thermosensitive ABA gel-forming triblock copolymers with controllable and tailorable properties. By varying polymer length and concentration the elastic properties of the hydrogels as well as their mesh size, swelling and erosion behaviour can be tuned. We show that the hydrogel networks are highly dense and that the decrease of gel volume is mainly the result of surface erosion, which in turn depends on both temperature and initial polymer concentration. In addition, we also show that the release kinetics of an entrapped protein is governed by a combined mechanism of erosion and diffusion. The prevalence of one or the other is strongly dependent on polymer concentration. Most importantly, the encapsulated protein was quantitatively released demonstrating that these hydrogels offer great potential as drug delivery systems.

4.1 INTRODUCTION

The potential of hydrogels as biomedical materials has long been recognised[1-3]. For this reason they are increasingly being utilised as delivery systems for a variety of therapeutic agents, especially because of their capacity to preserve the structure and functionality of incorporated drugs, particularly pharmaceutical proteins, and because they are usually well tolerated by living tissue[3, 4]. Hydrogels are hydrophilic polymeric matrices that have the ability to absorb large amounts of water and still maintain a distinct 3D structure[4, 5]. They are formed by either chemical (covalent) or physical (reversible) cross-links between hydrophilic polymer chains, which prevents these chains from dissolving in water. In recent years, there has been an increasing interest in physically cross-linked hydrogels, because the use of crosslinking agents can be avoided with these systems. These agents can not only affect the integrity of the substance to be entrapped (e.g. proteins, cells) but they are often toxic compounds which have to be removed/extracted before the gels can be applied in the body[5]. Because proteins possess many mechanisms for the physical association of peptide segments, they can be used to form gel networks in aqueous solution. Peptide-based biomaterials are attractive because of their programmable, biodegradable and bioresorbable nature[2]. A classic hydrogel-forming protein is gelatin, a natural polymer derived from collagen, which is a biocompatible and biodegradable material widely used for pharmaceutical and medical applications[6-8]. However, the variability in composition and structure of animal-derived gelatins, and the possibility of inflammatory responses in humans urged the development of techniques for the recombinant production of collagen and gelatin[6, 9]. A major advantage of recombinant production is the ability to genetically control the protein structure and thereby, its physical and chemical properties[5, 10]. Despite the developments in the past decade in the production of recombinant gelatins[11], their use as protein delivery systems is scarce[12-14].

In chapter 2 and 3 (see also reference[15]) we describe the genetic design, recombinant production and characterisation of a new class of ABA triblock copolymers forming thermosensitive gels with highly controllable and predictable properties. Gel formation is obtained by combining collagen-inspired (Pro-Gly-Pro)$_n$ end-blocks (T), which have triple helix-forming ability, with highly hydrophilic random coil blocks defining the distance between the trimer forming end-blocks. These monodisperse triblock copolymers have a defined molecular weight, and controllable physical-chemical properties. Importantly, they form gels with a molecular architecture that is much more defined than that of traditional gelatins.

We have shown that, small, but tailored changes in length of the different blocks result in significant changes in their properties (chapter 3). The thermostability of the hydrogels formed by these polymers can be tailored by changing the number of (Pro-Gly-Pro) repeats, and the viscoelastic properties can be changed by varying the polymer concentration and the length of the mid-block (chapter 3 and reference[15]). This makes these gel forming peptide-based triblock copolymers promising materials for drug delivery applications where the possibility of working with a predictable and tailorable material can be crucial for the success of a specific application.

The aim of the present work was to study the *in vitro* kinetics of gel erosion and protein release from this new type of hydrogels. Bovine Serum Albumin (BSA) was used as model protein for *in vitro* release studies. The hydrogel forming collagen-inspired triblock copolymers, denoted as TR4T and TR8T (chapter 2 and 3), were investigated. These consist of a mid-block made of 4 or 8 repeats of highly hydrophilic random coil R blocks and triple helix forming T end-blocks of 9 repeating units of (Pro-Gly-Pro) resulting in a gel melting temperature of ~37 °C (chapter 2 and 3, and reference[15]). Temperature, polymer concentration and the length of the random mid-blocks of the recombinant polymers were varied to establish the potential applicability of these materials as drug delivery matrices. Results were correlated with rheological studies, mesh size of the gel and gel degradation kinetics.

4.2 MATERIALS AND METHODS

Materials

Dextran Blue from Leuconostoc ssp. was obtained from Fluka (Buchs, Swuitzerland). Bovine serum albumin (BSA), Myoglobin, sodium azide (NaN_3) and trifluoroacetic acid (TFA) were provided by Sigma-Aldrich (Zwijndrecht, The Netherlands). Phosphate buffer saline pH 7.4 (PBS (8.29 g/l NaCl; 3.1 g/l $Na_2HPO_4.12H_2O$; 0.3 g/l $NaH_2PO_4.2H_2O$)) was purchased from Braun Melsungen AG (Germany) and acetonitrile (ACN) was purchased from Biosolve (Valkenswaard, The Netherlands).

Biosynthesis of triblock copolymers

The construction of the genes encoding TR4T (~42 kDa) and TR8T (~78 kDa), their transfection to *Pichia pastoris*, production and purification has been previously described in detail in chapter 2 and 3. The complete amino acid sequence of TR4T and TR8T has been deposited in GenBank[16] under the accession numbers ACF33479 and ACF33481, respectively.

Rheological characterisation

The rheological measurements where preformed with an Anton Paar Physica MCR301 rheometer, equipped with a cone and plate geometry of 50 mm diameter. Temperature control was insured by a Peltier system, which allowed fast heating and cooling. Polymer solutions were prepared in PBS buffer pH 7.4 containing 0.02 % (w/w) of NaN_3 (PBS/NaN_3) buffer, heated for half an hour at 60 °C and then introduced in the pre-heated geometry. The system was subsequently quenched to 20 °C to induce gel formation. A solvent trap was used to prevent evaporation. Gelation was followed for 5h at 20 °C by measuring the storage modulus (G’), as well as the loss moduli (G”) at a frequency of 1 Hz and a controlled strain of 1 %.

The average molecular weight between cross-links, $M_{c,}$ was calculated from the plateau storage modulus (G_0) using Equation (1)[15, 17], derived from the classical rubber-like theory and taking in consideration that only the trimer forming side blocks (T) form helices:

$$G_0 = \frac{FC_pRT}{M_c} \tag{4.1}$$

where C_p is the protein concentration (g/m^3), *R* the molar gas constant, T is the absolute temperature and F the front factor taken to be 0.5 as a trade off between the classical affinity network theory and the phantom network model[15].

Hydrogel swelling and degradation studies

The triblock copolymers were dissolved at a concentration of 10 % or 20 % (w/w) (total weight 60 mg) in PBS/NaN_3 buffer at 60 °C in cylindrical shaped glass vials with a diameter of 5 mm. Next 20 µl of a 100 mg/ml (w/w) solution of Dextran Blue (Mw=2,000,000 g/mol) in PBS/NaN_3 buffer was added to the polymer solutions and the mixture was heated to 42 °C for 10 min to allow homogeneous distribution of the Dextran Blue in the hydrogel. The gels (total weight approximately 100 mg) were allowed to stabilise overnight at room temperature. Next, the exact height of the gel inside the glass vials was measured (H_0) and 0.7 ml of PBS/NaN_3 buffer was added on the top of the gels and the vials were placed in a shaking water bath at 20 °C or 37 °C. At regular time intervals aliquots of 0.15 ml release buffer were taken, followed by replacement of the aliquot with fresh buffer, and measurement of the height of the gels (H_t) to determine the swelling ratio ($SR=H_t/H_{0;}$ only for 20 °C experiments). The removed buffer samples were analyzed for its Dextran Blue content by measuring the absorbance at 620 nm with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA).

Protein release studies

In vitro protein release from gels was studied using BSA (Mw_{BSA}=67,000 g/mol; d_{hBSA}=7.2 nm[18]) and myoglobin (Mw_{myo}=17,000 g/mol; d_{hmyo}=4.2 m[19]). Twenty µl of a concentrated solution of model protein (100 mg/ml) was added to 60 mg of a 10 % or 20 % (w/w) polymer solution in PBS/NaN_3 buffer, as described above in the hydrogel swelling and degradation studies section. The gels were allowed to stabilise overnight at room temperature. Next, 0.9 ml of PBS/NaN_3 buffer was added on the top of the gels and the vials were incubated in a shaking water bath at 20 °C or 37 °C. At regular time intervals aliquots of 0.15 ml release buffer were taken, and replaced by an equal amount of fresh buffer. The concentration of BSA or myoglobin in the samples was determined by using a Waters Acquity Ultra Performance LC system (UPLC, Waters, USA), with a BEH C18 1.7 µm, 2.1×50 mm column. An eluent gradient, from 0 to 100 % of eluent A was used, where eluent A was 95/5/0.1 % H_2O/acetonitrile/CF_3COOH and eluent B was 100/0.1 % acetonitrile/CF_3COOH. The injection volume of the samples was 5µl, the flow rate 0.25 ml/min and the detection of BSA was done at 210 nm and Myoglobin at 410 nm.

4.3 RESULTS AND DISCUSSION

Rheological properties

In this study, two triblock collagen-like proteins were used, further denoted as TR4T (~42 kDa) and TR8T (~78 kDa). These consisted of the same triple helix forming T end-block (see introduction) and a mid-block made of either 4 or 8 tandem repeats of a highly hydrophilic random coil R block (chapter 2). The proteins were produced and secreted in intact form by the yeast *P. pastoris* as host system (chapter 2 and 3). Their rheological properties have been extensively studied and reported previously (chapter 2, 3 and reference[15]). However, because the polymer concentrations used here are higher than used in the earlier rheological studies, we measured the storage (G') and loss modulus (G") of TR4T and TR8T at the working concentrations and determined the molecular weight between cross-links (M_c). Figure 4.1 shows the storage modulus of the different hydrogels as a function of time, measured at 20 °C for 10 and 20 % (w/w) formulations. The G' increased and reached a plateau after 3h or less, indicating that gel formation was complete. As expected, the storage G' increased with increasing polymer concentration. In addition, at the same gel concentration the G' was always higher for TR4T gels. These results are not surprising since, e.g., 20 % (w/w) TR4T gel contains a two times higher molar concentration of trimer forming end-blocks (T) than a 20 % (w/w) TR8T gel. Since the trimer forming end-blocks are responsible for trimer self-assembling, essential for the network formation, a higher molar ratio of T end-blocks results in a higher G'.

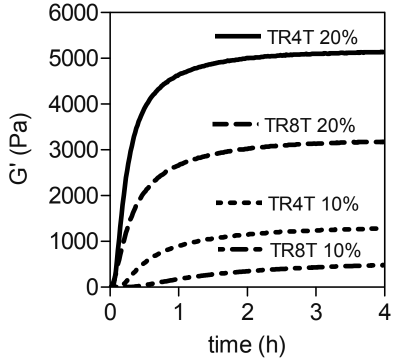

Figure 4.1 Characterisation of TR4T and TR8T 10 % and 20 % (w/w) hydrogels by dynamic rheology at 20 °C in time.

The measured plateau storage modulus (G_0) enabled the calculation of the, M_c, using Equation (4.1). The M_c values increased with decreasing polymer concentration and where higher for TR8T than for TR4T gels, at the same concentration (Table 4.1). Note that only for TR4T 20 % (w/w) gels the M_c value was below the molecular weight of the incorporated protein (BSA 67 kDa). In all other cases the M_c was higher than the molecular weight of the incorporated protein.

Table 4.1 Rheological data of TR4T and TR8T hydrogels at 20 °C

Polymer	Polymer Concentration (w/w)	G' (Pa)	M_c (kDa)
TR4T	10 %	1300 ± 7	94 ± 6
	20 %	5120 ± 10	47 ± 2
TR8T	10 %	548 ± 1	222 ± 3
	20 %	3180 ± 4	78 ± 1

Hydrogel stability

The stability of the hydrogels was investigated by measuring the time-dependent changes of their volume (reflected by their height). Figure 4.2 shows the swelling/dissolution profiles of TR4T and TR8T gels at a polymer concentration of 10 and 20 % (w/w). In line with their higher mechanical strength and thus the highest cross-link density, TR4T gels were stable for longer periods than TR8T gels at the same weight concentration. Figure 4.2 shows that at 20 °C, TR4T 20 % gels fully dissolved after 18 days, while TR8T gels lasted only 12 days. At higher polymer concentration, the hydrogels were stable for longer periods and reached a higher swelling ratio. When gels are incubated with buffer they absorb water, swelling stress accumulates and acts as an opposing force against the thermosensitive trimer formation that holds the gel together. When swelling stresses are high, dissociation of junctions occurs, eventually leading to polymer dissolution. Although the present hydrogels are able to resist the swelling stress to a certain extent, such phenomenon is highly dependent on temperature and surface erosion. At 20 °C both TR4T and TR8T 20 % gels swell to a maximum of ~2.8 times their initial volume, while 10 % gels only swell to a maximum of ~1.3 times their initial volume. The time to reach maximal swelling increases with concentration: in 120 h for 20 % gels and 48 h for 10 % gels.

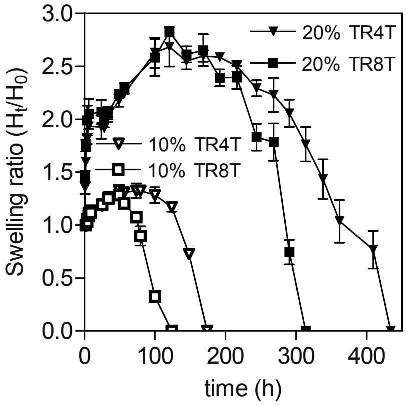

Figure 4.2 Swelling curves of TR4T and TR8T 10 % and 20 % (w/w) hydrogels at 20 °C as a function of time. Data are shown as mean ± standard deviation, n=3.

At 37 °C, close to the hydrogels melting temperature, hardly any swelling was observed. At this temperature, the dissociation rate of the trimer forming junctions and dissolution of the separate polymer chains is much faster than the rate of water uptake.

To evaluate the hydrogel degradation more accurately, at both 20 °C and 37 °C, release of Dextran Blue into the incubation buffer was monitored in time. Dextran Blue is a very large molecule, with a molecular weight of ~2 MDa, and is therefore essentially immobile in the network and can only be released by surface erosion. Figure 4.3a-b and Figure 4.4 show that Dextran Blue was released at a constant rate over time and 100 % release of the loaded amount was reached when the gels were fully dissolved. It also demonstrates that an increasing polymer concentration influences the release kinetics of Dextran Blue. For both TR4T and TR8T at both 10 and 20 % (w/w) polymer concentration, the release kinetics of Dextran Blue correlate well with the time dependent height variation shown in Figure 4.2, confirming that the degradation of the gel at 20 °C is mainly mediated by surface erosion. These observations can be confirmed by fitting the experimental Dextran Blue release data to the Ritger-Peppas equation (Equation 4. 2)[20]:

$$\frac{M_t}{M\infty} = kt^n \tag{4.2}$$

where M_t/M_∞ represents the fractional release of the entrapped compound, k is a kinetic constant, t is the release time and n is the diffusion exponent that can be related to the release mechanism of the entrapped molecules. If n=0.5 the release is controlled by Fickian diffusion, if n=1 the release is controlled by surface erosion, and if 0.5< n <1 then both diffusion and erosion play a role in the release mechanism. For the Dextran Blue release profiles at 20 °C (Figure 4.3a–b), the calculated n values were between 0.93 and 1.00. The release of Dextran Blue at 20 °C presents thus a zero order kinetics profile due to surface erosion of the hydrogel. As previously reported by van de Manakker et al. for cyclodextrin/cholesterol/PEG hydrogels[21], the assumption that Dextran Blue is trapped as immobile dye inside the network is correct and is therefore a reliable method to study hydrogel degradation.

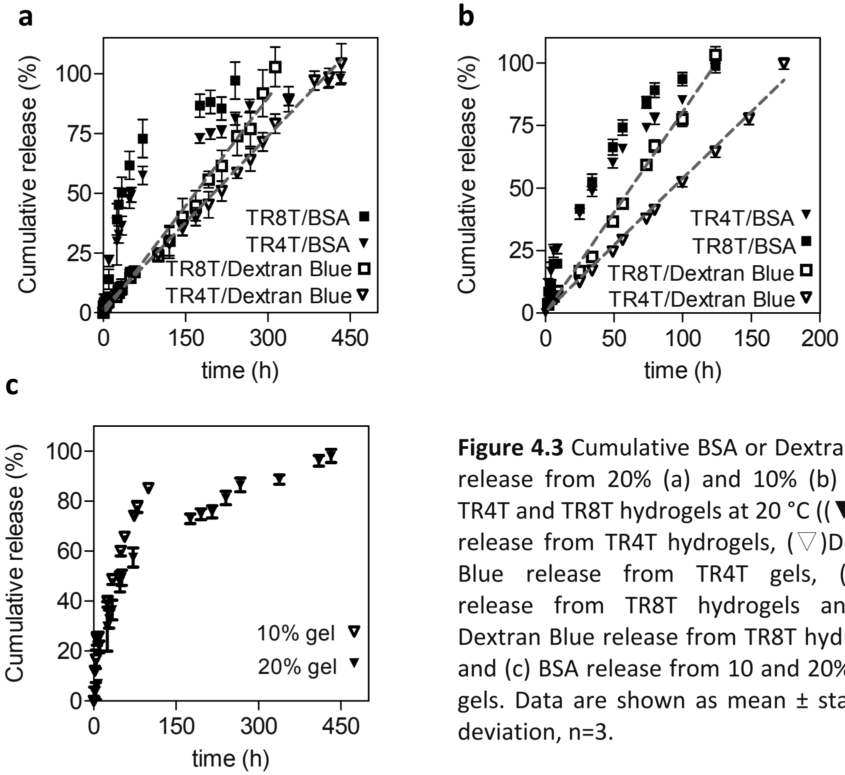

Figure 4.3 Cumulative BSA or Dextran Blue release from 20% (a) and 10% (b) (w/w) TR4T and TR8T hydrogels at 20 °C ((▼) BSA release from TR4T hydrogels, (▽)Dextran Blue release from TR4T gels, (■)BSA release from TR8T hydrogels and (□) Dextran Blue release from TR8T hydrogel); and (c) BSA release from 10 and 20% TR4T gels. Data are shown as mean ± standard deviation, n=3.

The release of Dextran Blue at 37 °C (Figure 4.4) presents also a zero order kinetics, but here, the hydrogel dissolution is accelerated by the higher temperature and complete release is reached already after 48 h, when the gel is fully dissolved. However, for applications at 37 °C a higher gel stability can be achieved by increasing the length of the terminal (Pro-Gly-Pro)$_n$ blocks.

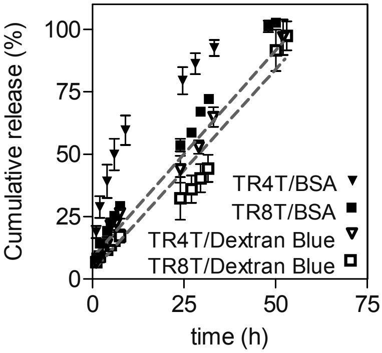

Figure 4.4 Cumulative protein release from 20 % (w/w) TR4T and TR8T hydrogels at 37 °C: (▼) BSA release from TR4T hydrogels, (▽)Dextran Blue release from TR4T gels, (■)BSA release from TR8T hydrogels and (□) Dextran Blue release from TR8T hydrogel. Data are shown as mean ± standard deviation, (n=3).

In vitro release studies

BSA (Mw 67 kDa) was used as model protein to study the influence of hydrogel composition and temperature on release. Figure 4.3-4.5 show that, at both 20 °C and 37 °C, BSA release from gels consisting of 10 and 20 % (w/w) TR4T or TR8T was sustained. Most importantly, quantitative (100 %) protein release was observed, meaning that BSA did neither aggregate nor precipitate. Also, protein release from all hydrogels at all polymer concentration is faster than the release of Dextran Blue which, as shown above, can only be released by surface erosion (Figure 4.3-4.4).

To characterise the release mechanism of BSA from the different hydrogels the experimental data were fitted to Equation (4.2). The diffusion exponents (n) vary between 0.5 and 1 showing that both erosion and diffusion play a role in the release kinetics (Table 4.2 and Figure 4.3a-b).

Table 4.2 Diffusion exponents (n) derived from entire set of BSA release data at 20°C.

Polymer	Polymer concentration (w/w)	Diffusion exponent (n)	Coefficient of determination (r^2)
TR4T	10 %	0.49 ± 0.03	0.99
	20 %	0.75 ± 0.07	0.97
TR8T	10 %	0.65 ± 0.08	0.99
	20 %	0.88 ± 0.09	0.98

For 10 % (w/w) TR4T the cumulative release scaled linearly with the square root of time up to a cumulative release of 90 % (inset in Figure 4.5a). These results point to a first order release kinetics[20], and thus, BSA release from 10 % (w/w) TR4T hydrogels is controlled by Fickian diffusion. At first sight, the calculated diffusion exponents (n) are higher for 20 % (w/w) gels, which would indicate that the release would be more surface- and less diffusion-controlled at these high polymer concentrations. However, a closer analysis of these results shows that after the first 24h, the cumulative release seems to be proportional to the square root of time and thus, to point to first order kinetics[20] (Figure 4.5b), indicating that protein release from the 20 % (w/w) gels only takes place after initial swelling of the gel. In the first hours, hardly any BSA is released from the gel, and release starts only after 10-24 h, when the swelling ratio is ~2, demonstrating that the networks are very tight and only when, sufficient water has been absorbed by the matrix the hydrogel mesh size is sufficiently large to allow protein diffusion. A similar sustained release profile was also found for other proteins proving the versatile character of these hydrogels as drug delivery system. The release kinetics for myoglobin can be found in Figure 4.6. Comparison of the release characteristics of the 10 and 20 % (w/w) gels shows that drug release from these hydrogels can be tailored by varying the recombinant polymer concentration (Figure 4.3c).

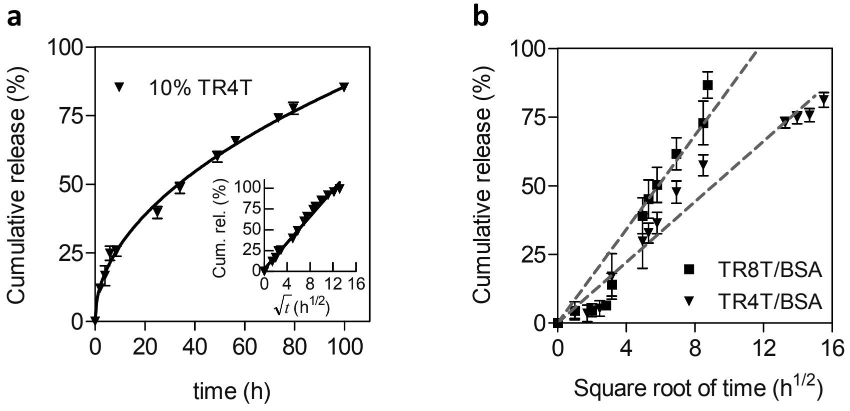

Figure 4.5 (a) Cumulative BSA release from 10 % (w/w) TR4T gels as a function of time and inset in function of square root of time. (b) Cumulative BSA release in order of the square root or time from 20 % (w/w) TR4T and TR8T hydrogels at 20 °C: (▼)from TR4T hydrogels, (■)from TR8T hydrogels.

At 37 °C, the release of BSA is largely determined by the melting rate of the hydrogels. In agreement with the higher equilibrium concentration of junctions (physical cross-links) remaining at this temperature in TR8T, as compared to TR4T (chapter 3)[22], the release of BSA (partly controlled by diffusion) and the gel erosion was somewhat slower in TR8T than in TR4T. Nevertheless, complete release from 20 % gels was reached after only 48 h even for TR8T gels, in agreement with the completion of the Dextran Blue release at the same polymer concentration and temperature (Figure 4.4). Using 10 % gels of TR4T or TR8T, complete release was even reached within 24 h (results not shown).

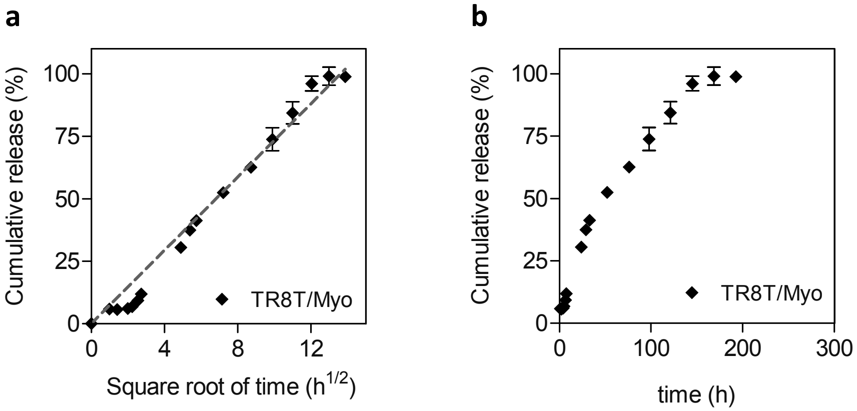

Figure 4.6 Cumulative Myoglobin release from 20 % (w/w) TR8T hydrogels at 20 °C (a) in order of time; and (b) in order of the square root of time. Data are shown as mean ± standard deviation (n=3).

4.4 CONCLUSIONS

The time-dependent release of entrapped protein from hydrogels consisting of a new class of telechelic triblock copolymers with collagen-like end-blocks was investigated *in vitro*, along with the time-dependent erosion of the gels. Apart from initial swelling, the volume of the gel gradually decreased due to surface erosion, which was governed by concentration and temperature. The gels showed a continuous and quantitative release of the entrapped protein by a combined mechanism of erosion and diffusion, the kinetics and relative predominance of which could be tailored by manipulating the concentration and design (i.e. the length) of the gelatin-like polymer. These hydrogels are attractive protein delivery systems for three reasons: (1) They did not appear to induce aggregation of entrapped protein and finally released 100 % of the entrapped protein in soluble form; (2) They are not chemically cross-linked, but self-assembling and self-healing[15, 23], and can be degraded into free amino acids without leaving a trace in the body; (3) Their design as telechelic triblocks automatically implies the freedom to tune their erosion and melting behaviour by varying the length and secondary structure of the mid-block and optionally, by designing end-blocks with different thermostability. The end-blocks used in the present polymers had a triple helix melting temperature of ~40-41 °C[15, 24] but for example, we have also developed end-blocks that melt at ~ 15-20 °C (unpublished data), or at ~57 °C[24]. Note that the gels melt when the percentage of end-blocks involved in trimolecular junctions becomes lower than fifty, which is always below the T_m of the triple helices[15, 22]. The present synthetic mid-blocks can also be replaced by human sequences, or by particular amino acid sequences that promote or avoid, depending on the purpose, the binding of specific organic molecules. The unique characteristics and the possibilities for tailored design of this material makes it a promising drug delivery system, but other pharmaceutical and biomedical applications are possible.

ACKNOWLEDGEMENT

We thank Roberta Censi for the help with the UPLC measurements and Paulina Skrzeszewska for the help with the rheological measurements.

This research was financially supported in part by The Netherlands Ministry of Economic Affairs and the B-BASIC partner organisations through B-BASIC, a public-private NWO-ACTS programme (ACTS=Advanced Chemical Technologies for Sustainability).

REFERENCES

1. S. R. Van Tomme, G. Storm and W. E. Hennink, *Int J Pharm*, 2008, **355**, 1-18.
2. S. Ramachandran and Y. B. Yu, *BioDrugs*, 2006, **20**, 263-269.
3. N. A. Peppas, P. Bures, W. Leobandung and H. Ichikawa, *Eur J Pharm Biopharm*, 2000, **50**, 27-46.
4. P. Gupta, K. Vermani and S. Garg, *Drug Discov Today*, 2002, **7**, 569-579.
5. W. E. Hennink and C. F. van Nostrum, *Adv Drug Deliv Rev*, 2002, **54**, 13-36.
6. D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen and J. Polarek, *Adv Drug Deliv Rev*, 2003, **55**, 1547-1567.
7. C. Yang, P. J. Hillas, J. A. Baez, M. Nokelainen, J. Balan, J. Tang, R. Spiro and J. W. Polarek, *BioDrugs*, 2004, **18**, 103-119.
8. S. Young, M. Wong, Y. Tabata and A. G. Mikos, *J Control Release*, 2005, **109**, 256-274.
9. M. Djabourov, ed., *Progress in Biotechnology*, Elsevier Science, Amsterdam, 2003.
10. A. A. Dinerman, J. Cappello, H. Ghandehari and S. W. Hoag, *Biomaterials*, 2002, **23**, 4203-4210.
11. J. Baez, D. Olsen and J. W. Polarek, *Appl Microbiol Biotechnol*, 2005, **69**, 245-252.
12. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
13. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
14. D. Olsen, J. Jiang, R. Chang, R. Duffy, M. Sakaguchi, S. Leigh, R. Lundgard, J. Ju, F. Buschman, V. Truong-Le, B. Pham and J. W. Polarek, *Protein Expr Purif*, 2005, **40**, 346-357.
15. P. J. Skrzeszewska, F. A. de Wolf, M. W. T. Werten, A. P. H. A. Moers, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2009, **5**, 2057-2062.
16. http://www.ncbi.nlm.nih.gov/genbank.
17. J. W. Goodwin and R. W. Hughes, *Rheology for Chemists: An introduction*, Royal Society of Chemistry, Cambridge, 2000.
18. E. W. Merrill, K. A. Dennison and C. Sung, *Biomaterials*, 1993, **14**, 1117-1126.
19. R. J. Stenekes, S. C. De Smedt, J. Demeester, G. Sun, Z. Zhang and W. E. Hennink, *Biomacromolecules*, 2000, **1**, 696-703.
20. P. L. Ritger and N. A. Peppas, *J. Controlled Release*, 1987, **5**, 37-42.
21. F. van de Manakker, K. Braeckmans, N. el Morabit, S. C. De Smedt, C. F. van Nostrum and W. E. Hennink, *Adv. Funct. Mater.*, 2009, **19**, 2992-3001.
22. H. Teles, P. J. Skrzeszewska, M. W. T. Werten, J. van der Gucht, G. Eggink and F. A. de Wolf, *submitted to Soft Matter*, 2010.
23. P. J. Skrzeszewska, J. Sprakel, F. A. De Wolf, R. Fokkink, M. A. Cohen Stuart and J. van der Gucht, *Macromolecules*, 2010, **43**, 3542–3548.
24. M. W. T. Werten, H. Teles, A. P. H. A. Moers, E. J. H. Wolbert, J. Sprakel, G. Eggink and F. A. de Wolf, *Biomacromolecules*, 2009, **10**, 1106-1113.

Chapter 5

Pilot-scale fermentation and purification of collagen-inspired triblock copolymers in *Pichia pastoris*

This chapter has been submitted for publication as: H. Teles, P. F. N. M. van Doeveren, F. A. de Wolf and G. Eggink.

SUMMARY

We have recently developed a new class of gel-forming collagen-like tri-block copolymers with biomedical relevance. These consist of highly hydrophilic random coil mid-blocks, in combination with proline rich trimer-forming end-blocks. With the aim of producing higher amounts of protein material for future applications and research we have carried out, at pilot-scale, the fermentation and purification of five of these proteins with molecular weights ranging from ~42 kDa to ~114 kDa. *Pichia pastoris* strains were grown in a 140 l bioreactor (100 l working volume) using a three-phase fermentation process. The fermentation culture reached high cell densities, 400-500 g/l (wet weight), and all proteins were efficiently expressed and secreted into the fermentation medium at a concentration of ~700-800 mg/l of cell free broth. The downstream processing principles elaborated previously at lab-scale were successfully adapted to the larger scale and resulted in 80-95 % recovery. The purified proteins (purity of at least 98 %) were intact and showed a similar performance to those obtained using lab-scale procedures. The good productivity and efficient DSP shown in this study provides a promising perspective towards a potential further scale-up to industrial production of these proteins.

5.1 INTRODUCTION

Over the past decades, collagen-like proteins have emerged as effective and widely used biomaterials in a range of different clinical applications[1]. The use of recombinant techniques for the production of collagen and gelatin (denaturated collagen) with defined composition and structure, and thus predictable properties, has opened the way for the development of novel materials. Furthermore, recombinant microbial production eliminates the concerns related to the quality, purity and predictability of animal derived collagen and gelatin and the risks of transmission of infectious agents[2, 3].

Several microbial[4-7] and non-microbial[8-13] sources have been explored for the production of recombinant collagen and gelatin. However, microbial approaches have advanced faster and shown to be more attractive for commercialisation than non-microbial approaches[2]. The methylotrophic yeast *P. pastoris* is one of the favoured hosts for the recombinant production of these proteins, especially because of its high levels of heterologous gene expression and secretion of low levels of endogenous proteins, hence facilitating purification[14, 15]. In addition, recombinant products derived from *P. pastoris* fermentations were shown to have excellent biocompatibility as determined in animal studies[16].

Successful recombinant production and purification at small-scale of hydroxylated[17, 18], and non-hydroxylated[19] human and synthetic collagen and gelatin has been achieved in *P. pastoris* by several groups, including our own[20, 21]. In chapter 2 and 3, we have reported on the recombinant production and characterisation of a new class of collagen-inspired tri-block copolymers with highly controllable and predictable gel-forming properties (chapter 2 and 3). Gel formation is obtained by combining collagen-inspired (Pro-Gly-Pro)$_9$ end-blocks (T), inspired by and behaving like natural collagen, with highly hydrophilic random coil P or R blocks defining the distance between the trimer forming T end-blocks. The polar ‘P’ block has a collagen-inspired designer sequence consisting of (Gly-Xaa-Yaa) repeats and the ‘R’ block has the same amino acid composition, but quasi random amino acid sequence.

This concept allows the production of custom made precision gels with potential for biomedical applications[22-25] (chapter 2). Indeed, these materials have already shown to be attractive drug delivery matrices (chapter 4).

So far, high level (1-3 g/l) secreted production of four of such triblock copolymers, consisting of 461 or 857 amino acids and denoted as TP4T, TP8T, TR4T and TR8T, has been realised with *P. pastoris* in small (3 l) fermenters (chapter 2 and 3). Also, simple and effective purification procedures have been successfully developed at laboratory scale (chapter 2 and 3). However, we feel that to deliver the amounts of material necessary for clinical applications, methods for large-scale production and purification of these proteins need to be investigated. Consequently, as a further development to our previous works, we now report on the pilot-scale production (140 l) and purification of five members of this new class of collagen-inspired triblock copolymers. In addition to the above mentioned proteins, we also fermented for the first time a 1253 amino acid long (~114 kDa) protein belonging to the same class, denoted as TP12T. The preliminary characterisation of the proteins is also investigated and compared to those produced using small-scale procedures.

5.2 MATERIALS AND METHODS

Expression vectors and *P. pastoris* transformations

The construction of the genes encoding TP4T and TR4T (chapter 2) (both ~42 kDa), TP8T and TR8T (chapter 3) (both ~78 kDa) and their transfection to *P. pastoris* has been described in detail in chapter 2 and 3. The previously described vectors pMTL23-P8T, pMTL23-P4 and pCR4-TOPO-T were used to construct the gene encoding TP12T (~112 kDa). Vector pMTL23-P8T was digested with *DraIII* (5´ to the P8 gene) and desphosphorylated and a *DraIII*/*Van91I*-digested P4 block was inserted to yield pMTL23-P12T. This vector was then digested with *DraIII* (5´ to the gene P12) and desphosphorylated, and a second *DraIII*/*Van91I*-digested T block was inserted to yield vector pMTL23-TP12T. TP12T fragment was then cloned into *P. pastoris* expression vector pPIC9 (Invitrogen) via XhoI/EcoRI. TP4T, TR4T, TP8T, TR8T and TP12T DNA sequences (and translated amino acid sequences) have been deposited in the GenBank under acession numbers EU834227-EU834231. The expression vectors were linearised with *Sal*I and transformed into the *his4* GS115[26] strain. Transformation and strain selection of transformants was as described previously[21].

Media composition

Minimal glycerol medium contained: 1 % glycerol, 1.34 % YNB and 0.4 mg/l biotine. Fermentation basal salts medium contained, per l: 26.7 ml H_3PO_4 85 %, 1.175 g $CaSO_4.2H_2O$, 18.2 g K_2SO_4, 14.9 g $MgSO_4.7H_2O$, 4.13 g KOH, 40.0 g glycerol. Trace elements contained (PTM_1), per l: 6.0 g/l $CuSO_4.5H_2O$, 0.08 g/l NaI, 3.0 g/l $MnSO_4.1H_2O$, 0.2 g/l $Na_2MoO_4.2H_2O$, 0.02 g/l H_3BO_3, 0.5 g/l $CoCl_2$, 20 g/l $ZnCl_2$, 65.0 g/l $FeSO_4.7H_2O$, 0.2 g/l biotin, 5 ml H_2SO_4

Pilot scale fermentation

P. pastoris strains were cultivated in a stainless steel reactor (Applikon, Schiedam, The Netherlands) with a total volume of 140 ls, 39.1 cm inner diameter and 114 cm height) and a working volume of 100 ls, and a d/D ratio of 0.358 (ratio of stirrer diameter to vessel diameter). Probes and sensors were connected to the bioreactor to measure temperature, pH, dissolved oxygen, foam and methanol. The fermentations were run as a 3 phase process as follows. First, the fermentation was initiated by inoculating the fermenter containing 50 ls of basal salts medium with 8.4 % [vol/vol] of a pre-culture grown in minimal glycerol medium for 24 h at 30 °C and 250 rpm in baffled flasks. The temperature was controlled at 30 °C, the dissolved oxygen (dO_2) was maintained at 20-30 % saturation and controlled by the proportional integral derivative (PID) cascade controller, which automatically adjusts the stirrer speed, the airflow rate (max 20 l/min) and the pressure (overpressure 0.08-0.8 bar). The pH was set at 3.0. Throughout the fermentation, ammonium hydroxide solution (30 %) was used to adjust the pH and compensate for the acidification brought about by the metabolic activity of the yeasts. The airflow was set at 0.16 l/(min·l) (i.e. per l of initial volume) and antifoam was delivered as required. PTM_1 trace elements were added (0.44 % of the initial medium volume) to the fermentation broth 2h after inoculation.

When the glycerol in the batch medium was consumed by the cells, as indicated by a sharp increase of dO_2 concentration (dO_2 spike), the glycerol fed-batch phase was initiated. This occurred approximately 25 to 31 h after the beginning of the fermentation by pumping a 50 % (v/v) glycerol-water feed, supplemented with 12 ml PTM trace elements per l of mixture. The initial feed rate was 17.7 ml of glycerol-water mixture per hour per l of initial medium and after 1 h the feed rate was linearly decreased from 17.7 to 0 ml/h in the course of 3 h. A cumulative amount of 35.4 ml per l of initial medium of glycerol-water mixture was added to the broth. At the beginning of the glycerol fed batch phase, the airflow was set to 4 l/min (per l of initial volume).

The methanol fed batch phase (induction phase) was started 1 h after the glycerol fed-batch had been initiated. First a fixed amount of methanol was added to the culture medium to a final concentration of 0.15 % (vol/vol) (per l of initial medium plus inoculum) in order to relate the signal of the methanol sensor to the methanol concentration in the fermentation broth. Subsequently, the methanol concentration in the broth was fixed at the set value by controlling the methanol pump with the help of the methanol sensor (Raven Biotech stand-alone Model 2.1). The air stream was enriched with 10 % pure oxygen (relative to saturation with air) and was controlled by the agitation speed (maximally 700 rpm) through the PID cascade controller. The addition of methanol automatically started when the cells began to consume methanol after consumption of the remaining glycerol. The fermentation was terminated ~ 95-110 h after inoculation, usually when the stationary phase was reached, as indicated by a strong decrease in base and methanol consumption.

Pilot scale purification

Previous to the harvest, the pH of the fermentation broth was increased to 8 by adding ammonium hydroxide, so as to allow precipitation of medium salts. During this process, the temperature was decreased to 20 °C, the pressure was released and de agitation was set to100 rpm. Cell removal was done by microfiltration using a cross-flow filtration unit equipped with a flat-plate microfilter (SmartFlow Technologies, Optisep 11000, 5 m^2, 0.22 µm) mounted in a stainless steel filter holder. The extracellular fluid was collected into a 1500 l stainless steel vessel previously heat-sterilised and cooled to 4 °C. The cell broth was concentrated to a maximum of 30 ls and the concentrated cell broth was washed (diafiltrated) several times with demineralised water using the same filter, while the cells were continuously pumped through the filter and back into the bioreactor,. The final cell-free broth was 300 ls. The secreted collagen-inspired polymer was then precipitated from the cell-free broth by adding ammonium sulphate up to 50 % saturation while gently stirring (60 rpm), followed by incubation overnight at 4 °C. The next day, the slurry with the precipitated product was microfiltrated using the mentioned cross-flow filtration unit equipped with a *fresh* flat-plate microfilter (SmartFlow Technologies, Optisep 11000, 5 m^2, 0.22 µm). The retentate with the protein flocks was concentrated approximately 10 times, subsequently washed (diafiltrated) two times with 50 l of water saturated with 50 % ammonium sulphate (to avoid protein dissolution in this step) and finally concentrated to a total volume of 30 ls and diluted by adding 250 l of demineralised water, so as to dissolve the polymer product. The temperature was increased to 45 °C for approximately 15 min to melt possible gel structures and the solution was stirred at 100 rpm. The temperature was then decreased to 20°C and the stirring continued until all the protein was dissolved. The permeate resulting from the ammonium sulphate precipitation steps was stored in cubic vessels of 1 m^3 and analysed for its protein content by SDS page before being discarded.

The above-described ammonium sulphate precipitation and microfiltration of the precipitated protein polymer was repeated once more, after which the final resuspended protein pellet was desalted by ultrafiltration (SmartFlow Technologies, Optisep 11000, 10 m^2, PES 5 kDa cut off). During this procedure the protein solution was diafiltrated with demineralised water until the conductivity of the permeate emerging from the filter unit was lower than 100 µS. A final 'dead-end' microfiltration with a disposable Pall filter (Supor Acropale 200, 0.2 µm) served to remove any possible remaining contaminations. The microfiltered product was lyophilised. In view of potential contaminations stemming from the ultrafiltration step, enhanced purity was ensured by carrying out a second purification of the concentrated product in the laboratory. This would not have been necessary if we would have disposed of a completely closed, sterile system. Thus, the protein was resuspended again after lyophilisation in 4 l of Mili-Q and precipitated twice by addition of dry ammonium sulphate to 50-60 % saturation, 20 min. centrifugation in a SLA-3000 rotor (Sorvall) at 10000 rpm, 4 °C and dissolution of the pellets in the same amount of milli-Q water. Finally, 40 % (vol/vol) ethanol was added to the resuspended pellet. After centrifugation in a Sorvall 3000 rotor at 10000 rpm, (4 °C) the pellet was discarded and ethanol was added to the supernatant to a final concentration of 80 % (vol/vol). The protein pellet obtained after centrifugation was air dried dissolved in milli-Q, microfiltered using a Pall filter (Supor Acropale 200, 0.2 µm) and lyophilised.

SDS-PAGE analyses and densitometry

Fermentation supernatants and pure products were analysed by SDS-page using the NuPAGE Novex system (Invitrogen), with 10 % Bis-Tris gels, MES SDS running buffer and SeeBlue Plus2 pre-stained molecular mass markers. Gels were stained in Coomassie SimplyBlue SafeStain (Invitrogen) and distained in MQ water. Densitometric quantification of pure protein standards, ranging from 2 µg to 25 µg micrograms, and culture supernatants was done with a Biorad GS-800 densitometer and analised using Quantity One computer software.

Mass spectrometry

MALDI-TOF was performed using an Ultraflex mass spectrometer (Bruker). Samples were prepared by the dried droplet method on a 600 µm AnchorChip target (Bruker), using 5 mg/ml 2,5-dihydroxyacetophenone, 1.5 mg/ml diammonium hydrogen citrate, 25 % (vol/vol) ethanol and 1 % (vol/vol) trifluoroacetic acid as matrix. Measurements (20 Hz) were made in the positive, linear mode, with the following parameters: ion source 1, 20000 V; ion source 2, 18450 V; lens, 5000 V; pulsed ion extraction, 550 ns. Protein Calibration Standard II (Bruker) was used for external calibration.

Rheology

An Anton Paar Physica MCR301 rheometer, equipped with a stainless steel CC17 Couette geometry, gap size: 0.71 mm, bob radius: 8.3 mm, and sample volume of 3 ml, was operated at an angular frequency of 1 Hz and a strain of 0.1 %. Protein solutions (3.2 mM in 0.2 M sodium phosphate pH 3.0) were heated to 45 °C, introduced in the geometry, and subsequently quenched to 20 °C to induce gel formation. Melting of the gel was studied by a stepwise increase (2.5 °C/step) of the temperature from 20 to 65 °C over a total time span of 3h.

5.3 RESULTS

Production of collagen-inspired triblock copolymers at 140 l scale

P. pastoris strains encoding the genes of five extracellular collagen-inspired triblock copolymers, i. e. TP4T, TP8T, TP12T, TR4T and TR8T, were grown in a 140 l (100 l working volume) bioreactor. All proteins have the same basic design consisting of triple helix-forming T end-blocks of 9 repeating units of (Pro-Gly-Pro) and a mid-block made of 4, 8 or 12 tandem repeats of highly hydrophilic random coil 'P' or 'R' blocks. The proteins molecular sizes ranged from ~42 kDa to ~114 kDa. All proteins have been previously produced in 3 l bioreactors.

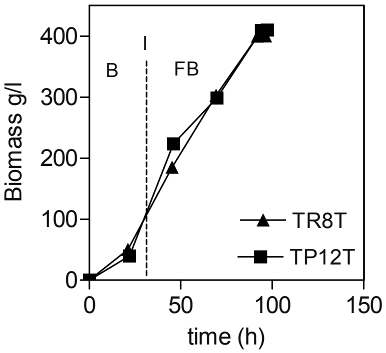

Figure 5.1 Profile of *P. pastoris* growth for TR8T and TP12T fermentations. Dotted line markes the moment of induction.

High cell density culture was scaled up using the three-phase growth strategy previously developed at laboratory (1-3 l) scale (chapter 2 and reference[27]). The whole fermentation process lasted for ~ 95-110 h and comprised a glycerol batch phase, a glycerol fed-batch phase and a methanol induction phase. The fed-batch cultivation mode was initiated ~ 25-31 h after inoculation by applying the glycerol-methanol mixed feeding strategy mentioned in Materials and Methods. Both biomass growth and methanol consumption were very similar between different fermentations. As an example, the time-dependent biomass formation (Figure 5.1) and methanol consumption (Figure 5.2) are shown for TR8T and TP12T fermentations. The cell wet weight (CWW) increased continuously and reached final values between 400 and 500 g/l (Table 5.1), and ~ 0.4 l of methanol was consumed per litre of fermentation broth. The growth rate during induction phase of both TR8T and TP12T fermentations was 0.021 h^{-1}. Similar growth rates were obtained for the other pilot scale fermentations. The accumulation of protein in the extracellular fluid of the fermentation broth was monitored by densitometric analysis of the product band in SDS-PAGE of samples of that cell-free fluid, in comparison to references of purified, freeze-dried product loaded on the same gel. Figure 5.3 shows the accumulation of protein in the culture supernatant at several time points for a TR8T fermentation. The accumulation of the target protein in the cell-free broth is indicated by a band of increasing intensity at the top of the gel. The protein concentration in the fermentation broth reached its maximum at the end of the fermentation. Clearly, these protein polymers show an anomalous migration behaviour in SDS-PAGE, as been documented in chapter 2 and 3.

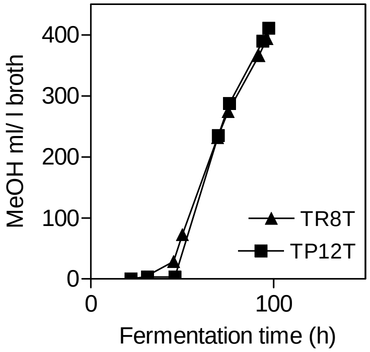

Figure 5.2 Methanol consumption profile for TR8T and TP12T fermentations

Secretion of recombinant TR8T could be detected after 69 h of fermentation and no traces of secreted protein are detected before methanol induction ($t_{ind.}$= 29h). The concentration of TR8T in the cell-free broth, determined accordingly, was ~700 mg/l. A similar value was estimated for TP4T, TP8T, TP12T and TP12T.

Table 5.1 Fermentation and purification data for TP4T, TP8T, TR4T, TR8T, TP12T

Protein	Theoretical size (kDa)	Cell wet weight (g/l)	Cultivation period (h)	Protein recovered (g)	Recovery (%)
TP4T	41.7	477	110	49	92
TP8T	78.2	498	98	53	93
TP12T	114.6	409	97	44	85
TR4T	41.7	410	100	53	94
TR8T	78.2	400	96	46	89

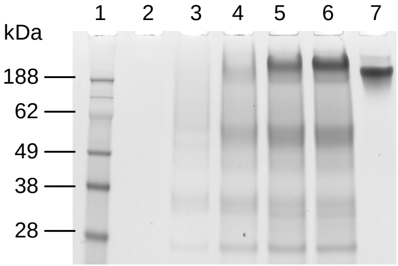

Figure 5.3 SDS-PAGE analysis of fermentation medium at several time points and purified TR8T. 1-Molecular mass marker, 2- t=25h (before induction), 3- t=45h, 4- t=69h, 5- t=91h, 6- t=96h (harvest), and 7- pure TR8T (10 µg). Lanes 1-5 all 15µl

Pilot scale purification of collagen-like protein polymers

Isolation of heterologous proteins secreted by *P. pastoris* is relatively simple as compared to other recombinant systems, because it grows on chemically defined media and does not secrete high amounts of endogenous protein[14, 15]. A clear supernatant can be obtained either by centrifugation or by filtration. We opted for cross-flow microfiltration. The protein was then purified from the cell-free supernatant by two steps of ammonium sulphate precipitation. The protein precipitate was separated from the resultant supernatant by cross-flow microfiltration and dissolved in demineralised water. Before lyophilisation, the product was desalted and concentrated by cross-flow ultrafiltration. Figure 5.4 shows a SDS-PAGE analysis of purified TP4T, TP8T, TP12T, TR4T and TR8T. The final recovery for each protein from the cell-free broth was 80-95 % (Table 5.1).

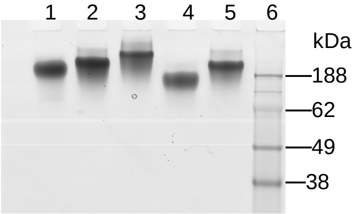

Figure 5.4 SDS-PAGE analysis of purified proteins. Lanes: 1-TP4T, 2- TP8T, 3- TP12T, 4- TR4T, 5- TR8T, 6- Molecular mass marker. Lanes 2-6: 20 µg of purified proteins.

Caracterisation

Amino acid analysis revealed a purity of at least 98 % of TP4T, TP8T, TP12T, TR4T and TR8T. The purified proteins were analysed by SDS-PAGE and all migrated as single bands indicating a high purity and intactness (Figure 5.4). MALDI-TOF MS analysis confirmed these results for the 42 kDa products (Figure 5.5) and 78 kDa products (not shown, with much more noise) and showed that the experimental values are in good agreement with the theoretical molecular weights (Table 5.1). It was not possible to obtain a clear MALDI-TOF MS spectrum of TP12T, because the molecular weight of this protein is close to the upper mass limit of the MALDI mass spectrometer used.

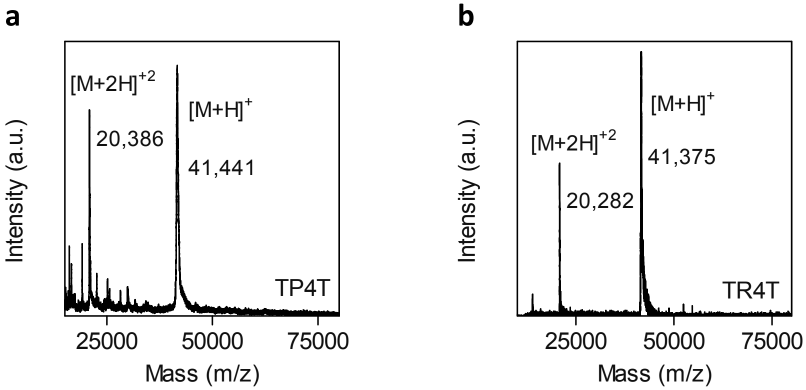

Figure 5.5 MALDI-TOF of purified (a) TP4T and (b) TR4T. Single and doubly charged molecular ions are indicated.

Gel formation by the proteins produced at pilot-scale was investigated by means of dynamic rheology. As an example, Figure 5.6 shows the development of the storage (G’) and loss (G’’) modulus for a 3.2 mM solution of a protein solution. As expected, at 20 °C, a physical gel is formed due to triple helix formation by the collagen-like T end-blocks. In the beginning of the gelation process viscous properties are predominant but within minutes, as the protein network develops, the elastic properties prevail. A steady-state (plateau) storage modulus is reached after roughly five hours. The protein melting behaviour was also studied by following the decrease of the storage modulus (G’) when the temperature was increased from 20 °C to 65 °C. The apparent melting temperature (T_m) of the gel can be deduced from the inflection point in the melting curve (inset of Figure 5.6) and was found to be ~ 37 °C. This value is in accordance with the T_m vales obtained for the same protein gel produced with lab-scale procedures (chapter 2). The proteins produced at pilot-scale thus appeared to be structurally intact and have similar properties to the proteins produced using lab-scale procedures.

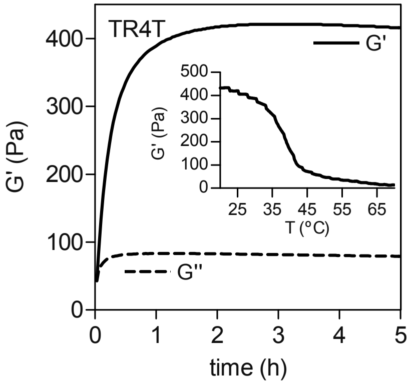

Figure 5.6 Characterisation of TR4T gel by dynamic rheology. Storage (G’) and loss modulus (G’’) as a function of time measured at 20 °C, 1Hz, γ= 1 %). The inset shows the storage modulus as a function of temperature

5.4 DISCUSSION

Collagen and gelatin are important materials used for several biomedical applications. But the variability in composition and structure from animal derived collagen and gelatin, and concerns related with the possibility that these materials might transmit infectious diseases, has accelerated the development of synthetic strategies for the production of this class of proteins. *P. pastoris* is the best studied and most used system for the high-yield production of animal or synthetic collagen and gelatine at laboratory scale[2, 28]. To produce the material necessary for clinical studies large amounts of product are required and efficient large-scale production strategies are needed. In this study, we meet these challenges by developing a pilot-scale strategy for the fermentation and purification of a new class of gel forming collagen-like proteins. Five different proteins were fermented and purified TP4T, TP8T, TP12T, TR4T and TR8T. The pilot-scale production of each of these proteins was initiated to foster high productivity of protein for future applications and research.

For the first time we carried out the fermentation culture in a 140 l fermenter and developed an efficient purification procedure for the production of large amounts of these proteins. High cell densities were reached (400<CWW<500 g/l of cell wet weight) and the expression level for all proteins fermented was ~ 700-800 mg/l of cell free broth. Although lower than the expression levels observed by us for lab-scale (3 l) fermentations (1-3 g of protein/l of cell free broth) (chapter 2 and 3), this level of expression is still above several other pilot-scale and large-scale fermentation processes using *P. pastoris* as a host system for the secreted production of heterologous proteins where secretion yields are below 500 mg/l[14, 29-32]. Also, the final biomass concentration obtained at pilot-scale reached slightly lower values than at small-scale (550-600 g/l cell wet weight). Cell growth is especially important for secreted proteins, as the concentration of product in the medium is roughly proportional to the concentration of cells in the culture[14, 15]. Thus, the lower yield obtained directly relates to the lower biomass obtained for the pilot-scale fermentations. The protein concentration in the cell-free broth of a small-scale TP4T fermentation with a final biomass (CWW) of 400 g/l was estimated to be ~ 600 mg protein per l of cell free broth. This value corresponds nicely with the protein concentration obtained for the pilot-scale fermentations at similar (final) biomass concentration. Dissolved oxygen is one of the most important factors for *P. pastoris* cell growth and heterologous protein expression. For this reason, the air is enriched with pure oxygen to increase productivity in *Pichia* fermentations[14, 15, 33]. In the lab-scale fermentations, we enriched the air going into the fermenter with 20 % pure oxygen. At pilot-scale it was not possible to enrich the air stream with the same percentage of pure oxygen. The bioreactor cooling system could not cope with the rise in temperature as a result of an increase in O_2 consumption. We had to decrease the percentage of pure oxygen in the air stream to ~10 % to avoid the temperature to rise above 30 °C, which would lead to cell death.

Harvest of products from *P. pastoris* fermentations can be done by centrifugation or filtration. At pilot scale, we chose for filtration rather than centrifugation, while we used centrifugation to obtain cell-free broth from the 3 l fermentations. Please note that filtration is always necessary for complete cell removal before purification. Centrifugation results in significant entrainment, making washing of the sediment a necessity for full recovery of secreted protein. For these reasons we opted for cross-flow microfiltration with open channel filters. The filtration system provided a high capacity and easy operation procedure. Several studies[34-36] have shown that cross-flow microfiltration is a reliable and cost-effective method for harvesting yeast cells at high cell density, with high recovery efficiencies. Also, scale-up of cross-flow microfiltration of high-cell density cultures is simpler to perform (linear scale-up) than other separation methods such as centrifugation[36].

Cross-flow microfiltration was an equally efficient and easy method to separate the precipitated target protein after ammonium sulphate precipitation. The total recovery from the cell free broth was approximately 80-95 %, higher than the recovery estimated for small-scale process (~30-40 %). The purified proteins were intact and showed similar performances to that observed for protein produced using lab-scale procedures.

In conclusion we have successfully developed a pilot-scale expression and purification scheme for the production of collagen-inspired triblock copolymers. Successful high-level expression and secretion was achieved, demonstrating the ability of *P. pastoris* to produce designer gelatins. The product was intact, had a high purity and was comparable to that produced using lab-scale procedures. The production scheme developed was found to be reproducible and easily scalable.

ACKNOWLEDGEMENT

We thank Marc Werten and Emil Wolbert for the construction, cloning, and transformation to *P. pastoris* of the TP12T gene; and Antoine Moers for testing the TP12T production, and purification method at small-scale.

This research was financially supported in part by the Netherlands Ministry of Economic Affairs and The B-Basic partner organisations (www.b-basic.nl) through B-Basic, a public-private NWO-ACTS programme (ACTS = Advanced Chemical Technologies for Sustainability).

REFERENCES

1. J. A. M. Ramshaw, Y. Y. Peng, V. Glattauer and J. A. Werkmeister, *J Mater Sci: Mater Med*, 2008, **20**, 3-8.
2. J. Baez, D. Olsen and J. W. Polarek, *Appl Microbiol Biotechnol*, 2005, **69**, 245-252.
3. W. Y. Aalbersberg, R. J. Hamer, P. Jasperse, H. H. J. Jongh, C. G. Kruif, P. Walstra and F. A. De Wolf, *Industrial proteins in perspective*, Elsevier, Amsterdam, 2003.
4. E. C. de Bruin, F. A. de Wolf and N. C. Laane, *Enzyme Microb Technol*, 2000, **26**, 640-644.
5. E. C. de Bruin, M. W. Werten, C. Laane and F. A. de Wolf, *FEMS Yeast Res*, 2002, **1**, 291-298.
6. D. D. Fan, Y. Luo, Y. Mi, X. X. Ma and L. Shang, *Biotechnol Lett*, 2005, **27**, 865-870.
7. P. R. Vaughn, M. Galanis, K. M. Richards, T. A. Tebb, J. A. Ramshaw and J. A. Werkmeister, *DNA Cell Biol*, 1998, **17**, 511-518.
8. M. Nokelainen, H. Tu, A. Vuorela, H. Notbohm, K. I. Kivirikko and J. Myllyharju, *Yeast*, 2001, **18**, 797-806.
9. L. Ala-Kokko, J. Hyland, C. Smith, K. I. Kivirikko, S. A. Jimenez and D. J. Prockop, *J Biol Chem*, 1991, **266**, 14175-14178.
10. A. Schnieke, M. Dziadek, J. Bateman, T. Mascara, K. Harbers, R. Gelinas and R. Jaenisch, *Proc Natl Acad Sci U S A*, 1987, **84**, 764-768.
11. M. Tomita, H. Munetsuna, T. Sato, T. Adachi, R. Hino, M. Hayashi, K. Shimizu, N. Nakamura, T. Tamura and K. Yoshizato, *Nat Biotechnol*, 2003, **21**, 52-56.
12. M. Tomita, N. Ohkura, M. Ito, T. Kato, P. M. Royce and T. Kitajima, *Biochem J*, 1995, **312 (Pt 3)**, 847-853.
13. D. C. John, R. Watson, A. J. Kind, A. R. Scott, K. E. Kadler and N. J. Bulleid, *Nat Biotechnol*, 1999, **17**, 385-389.
14. J. L. Cereghino and J. M. Cregg, *FEMS Microbiol Rev*, 2000, **24**, 45-66.
15. P. Li, A. Anumanthan, X. G. Gao, K. Ilangovan, V. V. Suzara, N. Duzgunes and V. Renugopalakrishnan, *Appl Biochem Biotechnol*, 2007, **142**, 105-124.
16. C. Yang, P. J. Hillas, J. A. Baez, M. Nokelainen, J. Balan, J. Tang, R. Spiro and J. W. Polarek, *BioDrugs*, 2004, **18**, 103-119.
17. O. Pakkanen, E. R. Hamalainen, K. I. Kivirikko and J. Myllyharju, *J Biol Chem*, 2003, **278**, 32478-32483.
18. O. Pakkanen, A. Pirskanen and J. Myllyharju, *J Biotechnol*, 2005.
19. D. Olsen, J. Jiang, R. Chang, R. Duffy, M. Sakaguchi, S. Leigh, R. Lundgard, J. Ju, F. Buschman, V. Truong-Le, B. Pham and J. W. Polarek, *Protein Expr Purif*, 2005, **40**, 346-357.
20. M. W. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
21. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
22. P. J. Skrzeszewska, J. Sprakel, F. A. De Wolf, R. Fokkink, M. A. Cohen Stuart and J. van der Gucht, *Macromolecules*, 2010, **43**, 3542–3548.

23. P. J. Skrzeszewska, F. A. De Wolf, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2010, **6**, 416-422.
24. M. W. T. Werten, H. Teles, A. P. H. A. Moers, E. J. H. Wolbert, J. Sprakel, G. Eggink and F. A. de Wolf, *Biomacromolecules*, 2009, **10**, 1106-1113.
25. P. J. Skrzeszewska, F. A. de Wolf, M. W. T. Werten, A. P. H. A. Moers, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2009, **5**, 2057-2062.
26. J. M. Cregg, K. J. Barringer, A. Y. Hessler and K. R. Madden, *Mol Cell Biol*, 1985, **5**, 3376-3385.
27. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
28. D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez, D. Carmichael, M. Perala, E. R. Hamalainen, M. Jarvinen and J. Polarek, *Adv Drug Deliv Rev*, 2003, **55**, 1547-1567.
29. P. Baumgartner, K. Harper, R. J. Raemaekers, A. Durieux, A. M. Gatehouse, H. V. Davies and M. A. Taylor, *Biotechnol Lett*, 2003, **25**, 1281-1285.
30. P. Baumgartner, R. J. Raemaekers, A. Durieux, A. Gatehouse, H. Davies and M. Taylor, *Protein Expr Purif*, 2002, **26**, 394-405.
31. N. Kong, X. Mu, H. Han and W. Yan, *Protein Expr Purif*, 2009, **63**, 134-139.
32. J. Q. Wang, F. Q. Yan, D. D. Wang, L. Ban, N. Sun, C. Y. Li, T. Zhang and W. Q. Yan, *Protein Expr Purif*, 2008, **60**, 127-131.
33. G. P. Cereghino, J. L. Cereghino, C. Ilgen and J. M. Cregg, *Curr Opin Biotechnol*, 2002, **13**, 329-332.
34. D. J. Bell and R. J. Davies, *Biotechnol Bioeng*, 1987, **29**, 1176-1178.
35. P. N. Patel, M. A. Mehaia and M. Cheryan, *J Biotechnol*, 1987, **5**, 1-16.
36. A. Wang, R. Lewus and A. S. Rathore, *Biotechnol Bioeng*, 2006, **94**, 91-104.

Chapter 6

General Discussion

Collagen and gelatin are the main structural functional components in many medical and pharmaceutical applications. Curently, the collagen and gelatin used in most of this applications is derived from animal sources, and their molecular structure and thermal properties cannot be controlled. In recent years, there have been efforts made to replace animal-derived collagen and gelatin by recombinant alternatives. Although much has been achieved, there are still several questions to be addressed.

The recombinant production of non-gel forming collagen-like proteins (gelatin) has advanced fast. Secreted production at high yields has been realised using *Pichia pastoris* as a recombinant system. As opposed, the production of recombinant collagen-like proteins with gel-forming capacity remains still a challenge. This is a drawback since more than 80 % of animal-derived gelatin is used in medical applications requiring gel formation, with the rest being applied mainly in injectable formulations (non-gelling gelatin)[1].

Irrespective of the expression system, recombinant gel-forming gelatin may be obtained by producing recombinant collagen and then denature it to yield recombinant gelatin. The production of gel-forming gelatin implies the use of a recombinant system capable of performing several post-translational modifications, the most important being the hydroxylation of proline residues into 4-hydroxyproline. The presence of hydroxyproline residues is important for the formation and thermostability of collagen-like triple-helices, but most of the microbial recombinant systems do not possess endogenous prolyl 4-hydroxylase (P4H), the enzyme performing the hydroxylation of proline in collagen chains. Nevertheless, successful production of fully hydroxylated triple-helical collagen could be achieved in *P. pastoris* by co-expression of heterologous P4H. However, the modest yields together with the complex downstream processing due to the intracellular accumulation of protein make this system unattractive for commercialization. Possibly due to similar reasons, other reported recombinant systems used to obtain hydroxylated triple-helical collagen have not advanced further. The low production yield of hydroxylated triple-helical collagen, combined with high process costs, might explain why there have been no reported attempts to convert recombinant collagen in recombinant gelatin. Because the production of correctly folded and processed triple-helical collagen in microbial systems is very complex, we have adopted in this thesis an alternative approach to obtain recombinant gel-forming gelatin without the need for hydroxylation. Gel formation is achieved due to the self-assembly of highly repetitive proline sequences present in artificial collagen-like polymers. In this way, some of the problems that are inherent to the production of recombinant hydroxylated collagen-like proteins can be avoided. Moreover, we show that these new protein polymers are an attractive alternative to animal-derived gel-forming gelatin as they seem to be suitable for biomedical or pharmaceutical applications.

Hydrogels from self-assembling collagen-inspired protein polymers

In **chapter 2** of this thesis, we discuss the design of a new class of non-hydroxylated gel-forming collagen-like proteins. The design of these proteins was based on the principle that chemically synthesized collagen-like peptides of the form $(\text{Gly-Pro-Pro})_n$ have the capacity to self-assemble in stable triple-helices. Although, by themselves, they are too short ($5< n <20$) to form gels, they can potentially be used in a multidomain system where they serve as junction points in a gel network. When designing protein polymers with gel-forming capacity two, apparently contradictory, behaviours should be taken into consideration: interchain interactions should be strong enough to keep the network together, and at the same time, the chains must allow solvent inside the network– otherwise the polymer will precipitate from solution rather than forming a swollen gel[2]. Gel-forming triblock protein polymers have been described before, making use of coiled-coils[2-7], elastin-mimetic motifs[8], or the fluorescent protein DsRed[6] as the cross-link-forming domain. In this thesis we describe the design (**chapter 2**) of a multidomain (triblock) class of proteins where the interchain binding domains consisting of $(\text{Gly-Pro-Pro})_n$ repeats (collagen-like domain) are spaced by a long, highly hydrophilic mid-block that assumes random coil structure in water (Figure 6.1). Because hydroxylation is absent, gel formation is only possible due to the high proline content of the end-blocks.

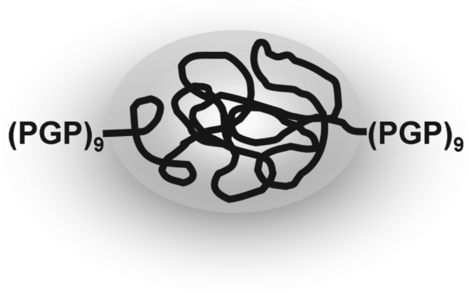

Figure 6.1 Single polypeptide consisting of a random-coil like mid-block flanked by collagen-like end-blocks (G-glycine, P-proline)

By changing the underlying DNA template, we designed several versions of these triblock copolymers. Those are made of 9 repeating units of (Pro-Gly-Pro) end-blocks (T), and mid-blocks made of tandem repeats of highly hydrophilic 9 kDa blocks (P_n or R_n) assuming random conformation: the polar 'P' block with a collagen-inspired designer sequence consisting of (Gly-Xaa-Yaa) repeats and the 'R' block with the same amino acid composition, but quasi-random amino acid sequence. Based on studies concerning helix melting of chemically synthesized $(\text{Pro-Pro-Gly})_n$[9-11], we tentatively choose nine (Pro-Gly-Pro) repeats (n=9) in order to provide a melting point in a biomedical relevant range.

Given the suitability of *P. pastoris* as a host organism to express and secrete recombinant collagen-like proteins, the DNA templates of several versions of triblock copolymers were inserted in the yeast genome. *P. pastoris* strains secreted several collagen-inspired triblock copolymers at high yields and in intact form (**chapter 2, 3 and 5**). Table 6.1 gives an overview of all the proteins fermented and their molecular weights.

Table 6.1 Collagen-inspired triblock copolymers fermented

Protein	TP4T	TP4T	TP8T	TR8T	TP12T	*TR12T
~Mw (kDa)	42		72		114	

*Strain available but never fermented

The same purification procedure, used previously for the non-gelling block protein P4[12], was successfully adapted for the purification of these proteins. The first two proteins analysed were TP4T and TR4T **(chapter 2)**. According to our expectations, these proteins formed thermoreversible hydrogels at room temperature. We have shown that the (Pro-Gly-Pro)$_9$ end-blocks in TP4T and TR4T are both exclusively and near-quantitatively involved in trimerization, as both the calorimetric enthalpies obtained with DSC and the van't Hoff enthalpies obtained with DSC and CD were in good agreement with values expected for free (Pro-Gly-Pro)$_9$[9-11]. Also the CD spectrum was consistent with the fact the ~88 % of the protein secondary structure is random coil and only ~12 % forms collagen-like helices. Since only the end-blocks exclusively form triple-helices, the molecular architecture of the gels is much more defined than that of animal-derived gelatins (Figure 6.2), where triple-helices can be formed along their entire chain. The cross-links in animal gelatin gels have a wide variety of lengths, compositions, and melting temperatures.

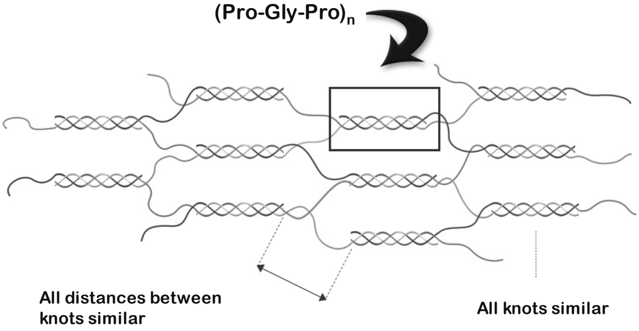

Figure 6.2 Schematic representation of hydrogels from collagen-inspired triblock copolymers

Gelling of the collagen-inspired triblock copolymers was, as opposed to animal gelatins, independent of the thermal history, meaning that the storage modulus reached a stable plateau within approximately 5 h, and did not increase significantly after the temperature was lowered. All polymers tested (**chapter 2, 3** and reference[13]) showed this same gelling behaviour. This difference between gels of collagen-inspired triblock copolymers and animal-derived is related with the fact that in the former all the helices have the same thermostability, while in animal-derived gelatins there are helices with different thermal stabilities and when the temperature is lowered additional helices of low thermal stability are formed, contributing additionally to the storage modulus.

Let us now look in more detail to the internal structure of these gels.

Hydrogel formation is driven by self-assembly of the T end-blocks that form triple-helices. Helices can be formed either by three T end-blocks coming from three different chains (junction) or by two of the three T end-blocks coming from the same molecule (loop). Only junction points with all three branches linked to the gel network connect elastically active chains, while loops and free (dangling) ends, do not[13] (Figure 6.3). According to classical gel theory, the storage modulus is proportional to the concentration of elastically active chains, i.e. chains that bridge two cross-links in the gel. The results obtained by Skrzeszewska et al[13] using TR4T, indicated that the fraction of chains that are active increases strongly with increasing concentration. The same relation was, not surprisingly, observed for TP4T, TP8T and TR8T (**chapter 3**). The probability that three end-blocks from three different chains form a junction point is higher than the probability of having a loop (or a free end), because at higher concentrations, the chains are closer together. A consequence of having fewer loops in the system is a higher number of active chains effectively contributing to the elastic properties of the network.

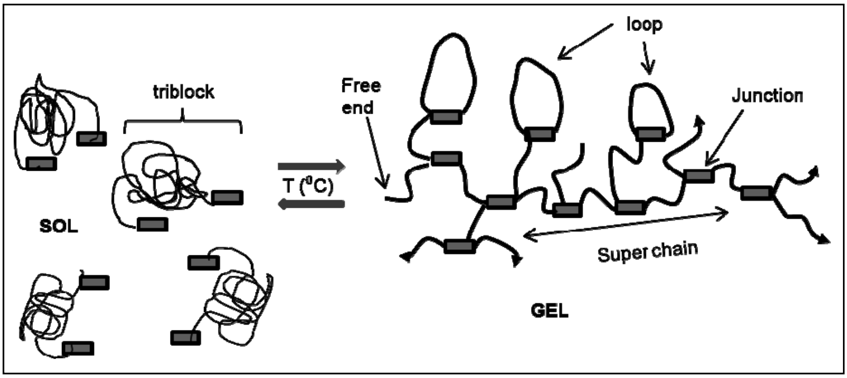

Figure 6.3 Network overview- Junctions, loops, and free (dangling) end

Protein concentration is also an important factor influencing the melting temperature of these gels. Note that it is the melting temperature of the gels and not the melting temperature of the helices which is influenced by polymer concentration. The triple-helices formed by the T end-blocks have a distinct melting temperature (T_m) that depends only on the number of (Pro-Gly-Pro) repeats. In the case of the present gels the helices melting temperature determined using differential scanning calorimetry and circular dichroism spectroscopy was ~41-42 °C (**chapter 2** and **3**). The T_m of the gels is lower than the T_m of the helices and varies with concentration[13]. The gels melt when the percentage of end-blocks involved in trimolecular junctions becomes lower than fifty percent, which is always below the T_m of the triple-helices (**chapter 2** and **3** and reference[13]), because the fraction of chain ends participating in junctions reaches fifty percent at significantly lower temperatures than the total helix content (junctions and loops). This effect is stronger for low polymer concentrations because, as explained above, the fraction of loops increases with decreasing concentration. These conclusions are based in a theoretical model developed by Skrzeszewska et al.[13].

We will see in the next section that temperature and concentration are not the only parameters affecting the physical-chemical properties of the network.

Minor structural changes result in dramatic changes in properties

The design of the synthetic collagen-like polymers described in **chapter 2** as telechelic triblocks implies the freedom to vary the length and secondary structure of the mid-block, and optionally, by designing end-blocks with different thermostability, tune the melting temperature of the hydrogels. This means that full control over the monomer sequence and polymer length allows the correlation between polymer structure and function, and, in this way, a tailored response of the hydrogels for a specific application may be achieved. This is an advantage as compared to animal-derived gelatin gels, where, due to the complexity of their composition, relatively little is known about gelatin structure and about structure-properties relationships.

In order to explore the relationship between the mid-block size and hydrogel-forming properties of these telechelic polypeptides with trimer forming collagen-like end-blocks four versions of this class of polymers, differing only in their mid-block length or amino acid sequence, were studied (**chapter 3**). For both long and short[13] versions, the storage modulus (measure for the elasticity of the gels) showed a strong dependency upon concentration and temperature. However, the longer versions of the collagen-like protein polymers, i.e., TP8T and TR8T had considerably higher storage modulus than their shorter counterparts, TP4T and TR4T, at the same molar concentration of trimer forming end-blocks. This difference is related to the fact that longer chains are less likely to form intramolecular loops, and therefore lead to a higher density of trimolecular network-building junctions at the same polymer concentration (**chapter 3** and reference[13]). Less expected was the difference in storage modulus observed between mid-blocks of the same length but different amino acid sequence, i.e., between TP4T versus TR4T, and TP8T versus TR8T. Although the P and R blocks are both random coils with exactly the same amino acid composition, their amino acid sequence is different- the R block is a ‘randomized version’ of the P block amino acid sequence. Fitzkee et al.[14] have shown that even a polypeptide chain that assumes a random coil conformation still has locally folded conformations that contribute to the overall flexibility of the chain. Apparently, this leads to a smaller radius of gyration for the P mid-blocks than for the R mid-blocks. This effect is stronger for the proteins with the shorter mid-blocks because the fraction of loops decreases with increased polymer length (**chapter 3**).

We have also observed that the length of the mid-block had an influence on the melting temperature of the gels (but not on the helices T_m). The temperature at which 50 % of the end-blocks were involved in junctions (T_m) was lower for the shorter polymers than for the longer polymers, as the loop fraction is higher for polymers with shorter mid-blocks. As a consequence, for the same molar concentration of end-blocks, the proteins with longer mid-bocks had higher T_m. The interpretation of the experimental results was supported by a model[14] developed earlier for the networks of these telechelic polymers with trimer-forming end-blocks. The model calculations correlated well with the experimental data showing that the well-defined (only one single type of cross-link is formed) multiplicity of the network of this particular class of hydrogels enables the prediction of the complete viscoelastic behaviour of the network when the polymer structure is well known. Our results suggest that, by controlling the structure of the present type of hydrogel-forming polymers their physical-chemical properties can not only be controlled and predicted but also changed in order to match a variety of different applications

In **chapter 3** we have only explored the effect of mid-block length and amino acid sequence in the viscoelastic behaviour of the networks. However, further research is necessary, to understand how the introduction of another type of block might affect the networks. For instance, the introduction of charged or hydrophobic blocks may change the intermolecular interactions between mid-blocks, by attraction or repulsion, and thus affect drastically the viscoelastic behaviour of the networks. In that scenario the present model would not be valid, new parameters would have to be included.

Applications

Numerous applications have been proposed and investigated for self-assembled hydrogels. There have been a wide-range of studies focused on the development of controlled release drug delivery systems using this material[15-20], especially because of the capacity of hydrogels to preserve the structure and functionality of incorporated drugs, particularly pharmaceutical proteins, and because they are usually well-tolerated by living tissue[21, 22].

It is generally accepted that one of the prerequisites for further progress in the design of drug delivery systems is the creation of macromolecular carriers with precisely defined structure and properties that match a physiological process[15]. It is for this reason that stimuli-responsive hydrogels are of significant interest. These hydrogels can undergo volume and phase transition induced by minor changes in the environment, such as temperature and pH. The present hydrogels from collagen-inspired triblock copolymers belong to such a class of materials because they undergo phase transition in response to temperature changes. As opposed to animal-derived gelatin, used often as a delivery system, they have a defined molecular weight and predictable physical-chemical properties (**chapter 2** and **3**).

So as to establish the potential applicability of these materials as drug delivery matrices, temperature, polymer concentration, and the length of the random coil mid-blocks of the recombinant polymers were varied (**chapter 4**). The time-dependent release of entrapped protein from TR4T and TR8T hydrogels was investigated *in vitro*, along with the time-dependent erosion of the gels. The gels showed a continuous and quantitative release of the entrapped protein by a combined mechanism of erosion and diffusion. The kinetics and relative predominance of which could be tailored by manipulating the concentration and design (i.e. the length) of the collagen-like polymer. The results suggest that these hydrogels are attractive protein delivery systems because they did not appear to induce aggregation of entrapped protein, and finally released 100 % of the entrapped protein in soluble form; and they are not chemically cross-linked, but self-assembling[13, 23], and can be degraded into free amino acids without leaving a trace in the body. In addition, their design as telechelic triblock implies that the erosion and melting behaviour can be tuned by varying the length and secondary structure of the mid-block and optionally, by designing end-blocks with different thermostability.

The hydrogels used in the study described in **chapter 4** have melting temperatures close to ~37 °C (**chapter 2** and **3**). At this temperature the gels erode rather quickly (1-2 days depending on the concentration used). For this reason the controlled release experiments were mainly preformed at 20 °C. At this temperature the hydrogels were stable over two weeks allowing the study of their erosion and controlled release kinetics. But controlled delivery applications require hydrogels stable at physiological temperature over extended periods of time (several weeks). The future use of this hydrogels as controlled release systems depends greatly upon the possibility of increasing their thermostability by increasing the number of Pro-Gly-Pro) repeats. However, there might be a limitation on the number of (Pro-Gly-Pro) repeats that can be added. Not only the amino acid sequence might become very repetitive leading to translational problems in the cell, which might result in low yields or no expression at all; but also end-blocks with higher thermostability (melting temperature) might induce trimer formation inside the yeast cell leading to secretion problems. The expression and production of collagen-inspired triblock copolymers with longer end-blocks is currently being addressed in another project.

Another important issue in developing controlled release systems for protein pharmaceuticals is the preservation of structure and activity of the incorporated protein, during and after casing. The incorporation of the model protein into the collagen-inspired hydrogels was done at ~42 °C. Below this temperature gel formation occurs almost instantaneously not allowing a uniform distribution of the model protein in the network. Many pharmaceutical proteins might loose their activity or structure at this temperature. Future *in vitro* studies are necessary to investigate the influence of the gel preparation procedure, as well as other factors such as mid-block characteristics (charge, isoelectric point, etc.), on the bioactivity of the incorporated drugs.

Production in *P. pastoris* and process scale-up

The development of efficient large-scale production processes can be a critical factor in whether or not a relevant pharmaceutical material is available in sufficient amounts to be used for application studies and eventually enter human clinical trials and the marketplace. In **chapter 5** we describe the development of a pilot-scale process for the fermentation and purification of this new class of gel forming collagen-like proteins. The procedures used for lab-scale fermentation (3 l bioreactors) and purification provided the basis for the development of a pilot-scale fermentation and purification scheme. *P. pastoris* strains were grown in a 140 litre bioreactor. The fermentation culture reached high cell densities, and all proteins fermented (proteins in table 6.1) were efficiently expressed and secreted into the fermentation medium at a concentration of ~700-800 mg/l of cell free broth. Although lower than the expression levels observed for lab-scale fermentations (1-3 g of protein/l of cell free broth) (**chapter 2** and **chapter 3**), this level of expression is still above several other pilot-scale and large-scale fermentation processes using *P. pastoris* as a host system for the secreted production of heterologous proteins, where secretion yields are below 500 mg/l[24-28].

The downstream processing used allowed a good recovery and protein with a high purity and similar performance to those obtained at lab-scale. Both fermentation and downstream processing processes and equipment chosen support linear scale-up under good manufacturing practices (GMP). Furthermore the use of *P. pastoris* is ideal for the production of proteins with pharmaceutical or medical relevance because the medium components are inexpensive and defined, and this yeast secretes low amounts of endogenous proteins facilitating purification of secreted recombinant proteins. The effort to generate certified production is low compared with other expression systems, where media components or by-products of the expression organisms can give rise to purification problems[29, 30].

The good productivity and efficient downstream processing shown in **chapter 5** provide a promising perspective towards a potential further scale-up to industrial production of these proteins. In fact, the pilot-scale scheme presented here could be readily translated to large-scale without need for further development. However, additional research may be necessary to improve yield, since this might be a bottleneck when producing enough material for future clinical applications and, ultimately, commercialization.

As pointed out in **chapter 5** a possible reason for the lower yield at pilot-scale as compared to lab-scale, was a decrease in dissolved oxygen inside the fermenter. Dissolved oxygen is often a limiting factor when a high growth rate is reached. This is even more critical in larger fermenters, because the oxygen can be depleted in some area of the reactor. Additionally, dissolved oxygen is one of the most important factors for *P. pastoris* cell growth and heterologous protein expression. For this reason, the air is enriched with pure oxygen to increase productivity in *Pichia* fermentations [27, 30, 31]. In the lab-scale fermentations, we enriched the air going into the fermenter with 20 % pure oxygen. At pilot-scale it was not possible to enrich the air stream with the same percentage of pure oxygen. The bioreactor cooling system could not cope with the rise in temperature as a result of an increase in O_2 consumption. We had to decrease the percentage of pure oxygen in the air stream to ~10 % to avoid the temperature to rise above 30 °C, which would lead to cell death. Further investigation should be done using a more efficient cooling system in order to be able to increase the percentage of pure oxygen in the air stream.

There are several other fermentation factors that can be modified in standard fermentation protocols to improve yield. Temperature and pH are two important fermentation parameters often influencing the yield of secreted proteins. Changes in standard protocols to both of these parameters in several *P. pastoris* fermentations have shown to positively influence yield and/or solubility of recombinant proteins in the fermentation medium[31-35]. We have carried out at lab-scale two independent fermentations where the pH and temperature were altered. The pH was set to 6 instead of 3[36], and the temperature to 20 °C[37] instead of 30 °C. In both cases there was no visible improvement in terms of protein yield. In particular the fermentation run at a lower temperature resulted in a significantly lower protein yield. From this it can be conclude that the pH and temperature conditions chosen for the fermentations are already optimum and no further improvements may be achieved by altering these factors.

Another way of optimizing protein expression in *P. pastoris* is by isolation of multicopy expression strains, i.e., strains that contain multiple integrated copies of the required gene[27, 38]. There are several studies were the increase in copy number of a gene could drastically enhance the expression levels of recombinant proteins[39-43]. However, it should be taken in consideration that there is an optimal copy number for a maximal protein production[43], and a systematic method should be applied to obtain efficient and stable high copy recombinant *P. pastoris* strains.

Final Conclusions and future prospects

Over the course of this project we designed, and produced successfully in yeast a novel class of gel-forming collagen-inspired triblock copolymers. Additionally, important steps have been taken in their characterization and potential application as drug delivery systems. The work described here demonstrates these gel-forming proteins are good candidates to replace animal-derived gelatin. As opposed to the latter, they have a defined molecular composition and predictable physical-chemical properties. Importantly, they form gels with a molecular architecture that is much more defined than that of traditional gelatins, allowing a thorough study of their structure and structure-function relationships.

Future studies should focus in exploring the tailorability of these protein polymers for a diverse range of pharmaceutical and medical applications. We have shown that these gels are good candidates for drug delivery. However, the introduction of different blocks with other characteristics and functionalities will create new possibilities and business opportunities. For example, the present hydrogels might be used as growth scaffolds for tissue formation, and, at the same time, act as a repository for various growth factors to facilitate tissue regeneration. The porosity of such scaffolds could be manipulated by varying polymer concentration and/or the length of the mid-blocks. Furthermore, the insertion of specific cell binding sequences in the mid-blocks could strongly enhance cell viability and proliferation in such scaffold, as has been observed for dextran-based hydrogels functionalised with certain peptide sequences[44]. Also, these hydrogels could be used to immobilize enzymes and cells, or be developed as *in situ* delivery systems for applications in cancer therapy. Another possibility is to insert mid-blocks with another type of folding pattern, or different stimulus responsiveness, e. g. pH responsive mid-block, which can lead to various functional materials.

For the further development of the present collagen-inspired self-assembling hydrogels as biomaterials, *in vivo* studies with therapeutic relevant proteins, and biocompatibility and cytotoxicity studies need to be performed. This would further enhance the potential application of these hydrogels in humans as controlled delivery systems or as scaffolds in tissue engineering applications. In particular, biocompatibility is a critical issue for the use of these materials in humans. One of the mid-blocks presented here showed favourable biocompatibility, as compared to animal products in blood applications[45], and the capacity to attract human cells in culture[46]. These results suggest these hydrogels are not expected to induce immunogenetic responses. In addition, their structural design as triblocks allows for example changes in the mid-block amino acid sequence to further enhance biocompatibility, e.g., by replacing it for a human amino acid sequence.

The development of new applications is more liable to be directed to the biomedical field because of their high added value and the need for safe and reliable alternatives to animal-derived collagen and gelatin. In the long run, as the microbial recombinant production develops and becomes progressively cheaper, even some technical and food applications might become feasible.

The contents of this thesis provide a good starting point for future development of this novel class of hydrogel forming collagen-like proteins.

REFERENCES

1. J. Baez, D. Olsen and J. W. Polarek, *Appl Microbiol Biotechnol*, 2005, **69**, 245-252.
2. W. A. Petka, J. L. Harden, K. P. McGrath, D. Wirtz and D. A. Tirrell, *Science*, 1998, **281**, 389-392.
3. Y. Cao and H. Li, *Chem Commun (Camb)*, 2008, 4144-4146.
4. L. Mi, S. Fischer, B. Chung, S. Sundelacruz and J. L. Harden, *Biomacromolecules*, 2006, **7**, 38-47.
5. C. Xu, V. Breedveld and J. Kopecek, *Biomacromolecules*, 2005, **6**, 1739-1749.
6. I. R. Wheeldon, S. C. Barton and S. Banta, *Biomacromolecules*, 2007, **8**, 2990-2994.
7. W. Shen, K. Zhang, J. A. Kornfield and D. A. Tirrell, *Nat Mater*, 2006, **5**, 153-158.
8. E. R. Wright, R. A. McMillan, A. Cooper, R. P. Apkarian and V. P. Conticello, *AdV. Funct. Mater.*, 2002, **12**, 149–154.
9. J. Engel, H.-T. Chen, D. J. Prockop and H. Klump, *Biopolymers* 1977, **16**, 601-622.
10. S. Frank, R. A. Kammerer, D. Mechling, T. Schulthess, R. Landwehr, J. Bann, Y. Guo, A. Lustig, H. P. Bachinger and J. Engel, *J Mol Biol*, 2001, **308**, 1081-1089.
11. K. Suto and H. Noda, *Biopolymers*, 1974, **13**, 2477-2488.
12. M. W. T. Werten, W. H. Wisselink, T. J. Jansen-van den Bosch, E. C. de Bruin and F. A. de Wolf, *Protein Eng*, 2001, **14**, 447-454.
13. P. J. Skrzeszewska, F. A. de Wolf, M. W. T. Werten, A. P. H. A. Moers, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2009, **5**, 2057-2062.
14. N. C. Fitzkee and G. D. Rose, *Proc Natl Acad Sci U S A*, 2004, **101**, 12497-12502.
15. J. Kopecek, *Eur J Pharm Sci*, 2003, **20**, 1-16.
16. S. R. Van Tomme, G. Storm and W. E. Hennink, *Int J Pharm*, 2008, **355**, 1-18.
17. M. Tang, R. Zhang, A. Bowyer, R. Eisenthal and J. Hubble, *Biotechnol Bioeng*, 2003, **82**, 47-53.
18. I. Y. Galaev and B. Mattiasson, *Trends Biotechnol*, 1999, **17**, 335-340.
19. Y. Qiu and K. Park, *Adv Drug Deliv Rev*, 2001, **53**, 321-339.
20. I. Roy and M. N. Gupta, *Chem Biol*, 2003, **10**, 1161-1171.
21. P. Gupta, K. Vermani and S. Garg, *Drug Discov Today*, 2002, **7**, 569-579.
22. N. A. Peppas, P. Bures, W. Leobandung and H. Ichikawa, *Eur J Pharm Biopharm*, 2000, **50**, 27-46.
23. P. J. Skrzeszewska, J. Sprakel, F. A. De Wolf, R. Fokkink, M. A. Cohen Stuart and J. van der Gucht, *Macromolecules*, 2010, **43**, 3542–3548.
24. P. Baumgartner, K. Harper, R. J. Raemaekers, A. Durieux, A. M. Gatehouse, H. V. Davies and M. A. Taylor, *Biotechnol Lett*, 2003, **25**, 1281-1285.
25. P. Baumgartner, R. J. Raemaekers, A. Durieux, A. Gatehouse, H. Davies and M. Taylor, *Protein Expr Purif*, 2002, **26**, 394-405.
26. N. Kong, X. Mu, H. Han and W. Yan, *Protein Expr Purif*, 2009, **63**, 134-139.
27. J. L. Cereghino and J. M. Cregg, *FEMS Microbiol Rev*, 2000, **24**, 45-66.
28. J. Q. Wang, F. Q. Yan, D. D. Wang, L. Ban, N. Sun, C. Y. Li, T. Zhang and W. Q. Yan, *Protein Expr Purif*, 2008, **60**, 127-131.
29. R. Fischer, J. Drossard, N. Emans, U. Commandeur and S. Hellwig, *Biotechnol Appl Biochem*, 1999, **30 (Pt 2)**, 117-120.

30. G. P. Cereghino, J. L. Cereghino, C. Ilgen and J. M. Cregg, *Curr Opin Biotechnol*, 2002, **13**, 329-332.
31. P. Li, A. Anumanthan, X. G. Gao, K. Ilangovan, V. V. Suzara, N. Duzgunes and V. Renugopalakrishnan, *Appl Biochem Biotechnol*, 2007, **142**, 105-124.
32. K. J. Potter, W. Zhang, L. A. Smith and M. M. Meagher, *Protein Expr Purif*, 2000, **19**, 393-402.
33. S. H. Wang, T. S. Yang, S. M. Lin, M. S. Tsai, S. C. Wu and S. J. Mao, *Protein Expr Purif*, 2002, **25**, 41-49.
34. V. Sarramegna, P. Demange, A. Milon and F. Talmont, *Protein Expr Purif*, 2002, **24**, 212-220.
35. M. M. Whittaker and J. W. Whittaker, *Protein Expr Purif*, 2000, **20**, 105-111.
36. C. I. Silva, *personal communication*.
37. H. Teles, *data not published*.
38. R. A. Brierley, in *Pichia pastoris protocols*, eds. H. D. R. and J. M. Cregg, Human Press, 2008, vol. 103, pp. 149-147.
39. M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek and F. A. de Wolf, *Yeast*, 1999, **15**, 1087-1096.
40. J. J. Clare, M. A. Romanos, F. B. Rayment, J. E. Rowedder, M. A. Smith, M. M. Payne, K. Sreekrishna and C. A. Henwood, *Gene*, 1991, **105**, 205-212.
41. M. Mansur, C. Cabello, L. Hernandez, J. Pais, L. Varas, J. Valdes, Y. Terrero, A. Hidalgo, L. Plana, V. Besada, L. Garcia, E. Lamazares, L. Castellanos and E. Martinez, *Biotechnol Lett*, 2005, **27**, 339-345.
42. A. Vassileva, D. A. Chugh, S. Swaminathan and N. Khanna, *Protein Expr Purif*, 2001, **21**, 71-80.
43. T. Zhu, M. Guo, Z. Tang, M. Zhang, Y. Zhuang, J. Chu and S. Zhang, *J Appl Microbiol*, 2009, **107**, 954-963.
44. L. S. Ferreira, S. Gerecht, J. Fuller, H. F. Shieh, G. Vunjak-Novakovic and R. Langer, *Biomaterials*, 2007, **28**, 2706-2717.
45. *U.S. Pat.*, 20050101531, 2005.
46. D. I. Rozkiewicz, Y. Kraan, M. W. T. Werten, F. A. de Wolf, V. Subramaniam, B. J. Ravoo and D. N. Reinhoudt, *Chemistry*, 2006, **12**, 6290-6297.

Summary

Collagens are the most abundant proteins in the animal and human body. Based on their structural role and compatibility within the body, collagen and gelatin (denaturated and partially degraded collagen) are a commonly used biomaterial in several medical and pharmaceutical applications. However, the variability in composition and structure of animal-derived collagen and gelatin presents a significant challenge for those using these proteins in medical applications. This is not the only concern related with the use of these materials in medicine. The potential presence of infectious agents, such as virus or prions, in collagen and gelatin derived from animal sources, and the possibility of inducing immunogenic reactions, poses a risk for patients receiving the medical product. In recent years, the biomedical relevance of collagen and gelatin and the advances in recombinant-based production systems have motivated scientists to find recombinant alternatives. Although much has been achieved there are still several questions to be addressed. The recombinant production of non-gel forming collagen-like proteins (gelatin) has advanced fast. Secreted production at high yields has been achieved using *Pichia pastoris* as a recombinant system. On the contrary, the production of recombinant collagen-like proteins with gel-forming capacity remains still a challenge. This is a drawback since gel formation is required in many medical applications using these proteins as functional component.

Recombinant (microbial) production of gel-forming collagen-like proteins presents various difficulties, not least because of the many post-translational modifications required for a recombinantly produced collagen molecule to achieve a fully folded, triple-helical conformation. The basis of the thermoreversible knots in gelatin is the characteristic triple-helical structure of collagen. It consists of repetitive (Gly-Xaa-Yaa) triplets, commonly containing approximately 22 % proline. At this low proline content, triple helices are not thermally stable above 5-15 °C unless, as occurs in animals, the prolines in the Yaa position are post-translationally modified to 4-hydroxyprolines by the enzyme prolyl-4-hydroxylase (P4H). Because microbial hosts generally lack this enzyme, recombinant production of thermally stable triple-helical collagen and gelatin requires coexpression of mammalian P4H.

The aim of this thesis is the design and microbial production of collagen-inspired designer polymers, forming hydrogels without the need for prolyl-hydroxylation, and with pre-defined and tunable physical-chemical properties. The yeast *P. pastoris* is used as microbial factory for the production of these proteins. In addition, we show that these new protein polymers are an attractive option to animal-derived gelatin as they appear to be potentially suitable for biomedical or pharmaceutical applications.

In **chapter 2** we describe the genetic design, recombinant production and preliminary characterisation of a new class of ABA triblock copolymers forming thermosensitive gels with highly controllable and predictable properties. Gel formation is obtained by combining proline-rich collagen-inspired (Pro-Gly-Pro)$_9$ end-blocks (T), which have triple helix-forming ability, with highly hydrophilic random coil blocks (P_n or R_n) defining the distance between the trimer forming end-blocks. We report the secreted production in yeast at several g/l of two such non-hydroxylated ~42 kDa triblock copolymers, TP4T and TR4T. The dynamic elasticity (storage modulus) of the gels from these collagen-inspired triblock copolymers was comparable to animal gelatin with a similar content of triple helices. In favourable contrast to traditional gelatin, the dynamic elasticity of the new material, in which only one single (well-defined) type of cross links is formed, is independent of the thermal history of the gel. The novel hydrogels have a ~37 °C melting temperature. However, the thermostability of the hydrogels formed by these polymers can be tailored by changing the number of (Pro-Gly-Pro) repeats. The concept allows to produce custom-made precision gels for biomedical applications.

In **chapter 3** it was shown that small, but tailored changes in the length of the mid-block of the collagen-inspired triblock copolymers results in significant changes in the viscoelastic properties of the hydrogels. We compared 4 different triblock copolymers, differing only in their mid-block size or mid-block amino sequence. The shorter versions, i.e. TP4T and TR4T, had mid-blocks made of ~400 amino acids, and their longer counterparts, i.e. TP8T and TR8T, ~800 amino acids. These results obtained indicate that the elastic properties of the network are not only a function of concentration and temperature but also of polymer length. The experimental results were well described by an analytical model that was based on classical gel theory and accounted for the particular molecular structure of the gels, and the presence of loops and dangling ends. These results suggest that, by controlling the structure of the present type of hydrogel-forming polymers through genetic engineering their physical-chemical properties can be predicted, and tailored in order to match a specific application

In **chapter 4** we explored the potential of hydrogels from collagen-inspired triblock co-polymers as drug delivery systems. We studied the erosion and protein release kinetics of two of these hydrogel-forming polymers, i.e. TR4T and TR8T, differing only in their mid-block length (mid-block molecular weights ~37 kDa and ~73 kDa). By varying polymer length and concentration, the elastic properties of the hydrogels as well as their mesh size, swelling and erosion behaviour can be tuned. We show that the hydrogel networks are highly dense and that the decrease of gel volume is mainly the result of surface erosion, which in turn depends on both temperature and initial polymer concentration. In addition, we show that the release kinetics of an entrapped protein is governed by a combined mechanism of erosion and diffusion. The prevalence of one or the other is strongly dependent on polymer concentration. Most importantly, the encapsulated protein was quantitatively released demonstrating that these hydrogels offer great potential as drug delivery systems.

The development of efficient large-scale production processes can be a critical factor in whether or not a relevant pharmaceutical material is available in sufficient amounts to be used for application studies and eventually enter human clinical trials and the marketplace. In **chapter 5** we describe the development of a pilot-scale process for the fermentation and purification of five collagen-inspired triblock copolymers (TP4T, TR4T, TP8T, TR8T and TP12T) with molecular weights ranging from ~42 kDa to ~114 kDa. *P. pastoris* strains were grown in a 140 liter bioreactor using a three-phase fermentation process. The fermentation culture reached high cell densities, and all proteins were efficiently expressed and secreted into the fermentation medium at a concentration of ~700-800 mg/l of cell free broth. The downstream processing principles elaborated previously at lab-scale were successfully adapted to the larger scale and resulted in 80-95 % recovery. The purified proteins were intact and showed a similar performance to those obtained using lab-scale procedures. The good productivity and efficient downstream processing (DSP) shown in this study provides a promising perspective towards a potential further scale-up to industrial production of these proteins.

In **chapter 6** some of the results obtained in the thesis are highlighted and suggestions for further research are given.

The contents of this thesis provide a good starting point for future development of this novel class of hydrogel forming collagen-like proteins.

Acknowledgements

When I decided to undertake a PhD I was far from imagining that I would learn as much at a scientific level as at a personal level. Unfortunately, this thesis does not summarise all the things I have learnt, both professionally and at a personal level. Doing a PhD is much like running your own company; its success depends on knowing when to take risks, making the right choices, listening to the right advice.
Along the way, there were a number of people that gave an enormous contribution, and without whom this thesis would not have been possible.

First, to my parents who despite feeling sad that I was so far from home gave me their support and encouragement. *Obrigado pelo amor que me dão, por me encorajarem e acreditarem sempre em mim*. To my sister who has unknowingly taught me to be more patient and less hard with myself.

My sincere acknowledgements to my co-supervisor, Frits de Wolf, and my promoter, Gerrit Eggink. Frits, thank you for your scientific insight, your bright ideas, and for your critical corrections and suggestions. You helped shape much of this thesis. Gerrit, thank you for giving me so much freedom during all these years, and for bringing things into a broader perspective – which is crucial for the success of a PhD thesis.

I am deeply indebted to Hans Tramper. Without your persistence, guidance and experience, I doubt I could have reached the end. Thank you so much for your unconditional support and for setting things back on track.

To Paulina for her invaluable help with the rheological measurements and to Jasper van der Gucht for all his expertise, critical corrections and for being available for all my questions and doubts.

I would like to express my gratitude to Wim Hennink and Tina Vermonden. It was a pleasure to work with you. Thank you for making me feel so welcome, and above all, thank you for your support, guidance and supervision during my stay at the Pharmacy

Department of Utrecht University. In just a few months I have learned a lot from both of you.
A big thank you to Roberta Censi for her support and help with the UPLC measurements. Thank you for our talks and your friendship. I also cannot forget all the other people I have met during my stay at the Pharmacy Department that made those four months so enjoyable.

To my first office mate, Louise, who has always been so sweet to me. To Catarina and David with whom I shared so many funny moments in our office; definitely the last four years would have been much duller if you were not around. The best office mates ever! Thank you both also for your friendship. David, another wonderful friend I made in the last years, *gracias* for your 'Spanish lessons', and your funny remarks; nobody will ever imitate me as well as you do, you give me the opportunity to have a good laugh about myself. Catarina, a friend from many years and many years to come, a colleague with whom I had the best scientific discussions, we definitely complement each other in making a great working team.

To all the people from lab 3.10/3.11 especially Jan, Emil, Marc and Twan, for their valuable help and technical support, always ready to answer to my questions. I would also like to give thanks to Arnoud and Kees for their help with gel filtration measurements.

To all the PhD students and all the people from the bioconversion group of Food and Biobased research for the *gezellig* environment and for sharing all the ups and downs of working in science.

My thanks to the MSc students Martijn and Pei for their valuable contribution to my thesis. A special thank you to Pei who also became a very good friend.

I also cannot forget Marloes, Martin, Wim, Teresa, Cristina and Eva, the lovely Dutch/Portuguese/Spanish lunch group that made the first year lunches so enjoyable. The lunches were never the same without you all.

Daniel, thank you so much for the corrections of parts of this thesis. Every time I learn more with you about this apparently easy, but after all challenging, language that is English.

To Cindy, Paula, and Manuel, for taking care of my body, my spirit and my soul. Thank you for your guidance. A big hug to the three. *Um abraço especial a ti Manuel que tiveste as palavras certas no momento certo.*

Laura, Leila, and Hagit, thank you for your unconditional friendship, for all the nice talks and fun moments. Laura and Leila: *minhas queridas amigas, é de facto fantástico que tivessemos acabado todas ao mesmo tempo na Holanda, eu teria sentido que algo faltava se vocês não tivessem estado por perto*; and Hagit: I am really happy I have met you, you definitely became a friend for life.

Last, but definitely not least, to my *liefje* Patrick without whom the achievements of the last five years would probably have been possible, but without whom everything would have been much more difficult, and less meaningful. Thank you for the bottom of my heart for all your support and help, thank you for giving me so much. You were the one that has always been there, always without exception, in the joyful moments, as well and in the difficult ones.

Muito obrigado a todos!
Dank jullie wel allemaal!

Helena

About the Author

CURRICULUM VITAE

Helena Maria de Albuquerque Ferreira Teles was born on the 19th February 1981 in Tondela, Portugal. She attended the local primary and secondary school. Between 1992 and 1998 she also studied Piano playing, Musical Formation and Music History at the Music Regional Conservatory Dr. José de Azeredo Perdigão in Viseu, Portugal. In 1998 she finished her secondary education and started studying Biological Engineering at the Technical University of Lisbon- Instituto Superior Técnico. Her graduation thesis on the study of a continuous process for the production of Maillard compounds and the thermodynamic analyses of the Maillard reaction, was conducted at the University of Reading- School of Food Biosciences, U.K.. After her graduation at the end of 2004 she worked for 2 more months on a follow-up of her graduation project at the University of Reading- School of Food Biosciences. In August 2005 she started her PhD at the Institute of Food and Biobased Research in collaboration with the Bioprocess Engineering group of Wageningen University, under the supervision of Prof. dr. G. Eggink and Dr. F. A. de Wolf. The research was focused on the production in yeast and functional characterization of a new class of collagen-like protein polymers. The results of this research are described in this thesis.

LIST OF PUBLICATIONS

M. W. T. Werten, H. Teles, A. P. H. A. Moers, E. J. H. Wolbert, J. Sprakel, G. Eggink, and F. A. de Wolf, "Precision gels from collagen-inspired triblock copolymers", *Biomacromolecules*, 2009, **10**, 1106-1113.

H. Teles, P. J. Skrzeszewska, M. W. T. Werten, J. van der Gucht, G. Eggink, and F. A. de Wolf, "Influence of molecular size on gel-forming properties of telechelic collagen-inspired polymers" *Soft Matter*, 2010, DOI:10.1039/C0SM00175A.

H. Teles, T. Vermonden, G. Eggink, W.E. Hennink, F.A. de Wolf, "Hydrogels of self-assembling collagen-inspired telechelic triblock copolymers for the sustained release of proteins'', *Journal of Controlled Release*, 2010. *Accepted for publication.*

H. Teles, P. F. N. M. van Doeveren, F. A. de Wolf and G. Eggink, "Pilot-scale production and purification of collagen-inspired triblock copolymers". *Submitted for publication*.

C. Silva, H. Teles, G. Eggink, F. A. de Wolf, M. W. T. Werten, "Influence of trimer stability in the production and secretion of collagen-inspired triblock copolymers in *Pichia pastoris*''. *In preparation*.

EDUCATIONAL ACTIVITIES

Discipline specific activities

Courses and workshops

Mass Spectrometry, Proteomics and Protein Technology (2006)
Advance course in *Pichia pastoris* Technology (2007)
DSM Workshop (2007)
Bioreactor course (2008)
Advanced course on Biocatalysis (2008)

Conferences and Symposiums

B-Basic symposium (2005, 2006, 2007, 2008, 2009)
Netherlands Biotechnology Conference (2006)
Protein expression Europe (2008)

General courses

VLAG PhD week (2006)
Presentation Techniques (2007)
Scientific Writing (2009)
Career Perspectives (2009)

Optional courses and activities

Brainstorm week Process Engineering (2005)
PhD trip Process Engineering (2006, 2008)
Genomics Researchers Event (2007)

The research described in this thesis was financially supported by the Netherlands Ministry of Economic Affairs and the B-Basic partner organizations (www.b-basic.nl) through B-Basic, a public-private NWO-ACTS programme (ACTS= Advanced Chemical Technologies for Sustainability).

Printed by GVO drukkers & vormgevers B.V. / Ponsen & Looijen, Ede.

Cover design by Patrick van Doeveren and Helena Teles