

Silk-Collagen-like Block Copolymers with Charged Blocks

Self-assembly into nanosized ribbons and macroscopic gels

Promotoren:

Prof. Dr. M.A. Cohen Stuart

Hoogleraar Fysische chemie met bijzondere aandacht voor colloïdchemie

Wageningen Universiteit

Prof. Dr. G. Eggink

Hoogleraar Industriële Biotechnologie, Sectie proceskunde

Wageningen Universiteit

Copromotor:

Dr. F.A. de Wolf

Senior wetenschappelijk onderzoeker

AFSG, Bioconversion

Wageningen UR

Promotiecommissie:

Prof. Dr. J. van der Oost

Wageningen Universiteit, Nederland

Prof. Dr. G.T. Robillard

Universiteit van Groningen, Nederland

Prof. Dr. R.J. Koopmans

Dow Europe, Freienbach, Zwitserland

Dr. Ir. C.P.M. van Mierlo

Wageningen Universiteit, Nederland

Dit onderzoek is uitgevoerd binnen de onderzoekschool VLAG

Silk-Collagen-like Block Copolymers with Charged Blocks

Self-assembly into nanosized ribbons and macroscopic gels

Aernout Anders Martens

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M.J. Kropff,
in het openbaar te verdedigen
op woensdag 10 september 2008
des namiddags te half twee in de Aula

Aernout A. Martens, Silk-collagen-like block copolymers with charged blocks, self-assembly into nanosized ribbons and macroscopic gels (2008)

PhD thesis, Wageningen university, Wageningen, The Netherlands – 160 pages

ISBN: 978-90-8504-976-0

To my family

Table of contents

	Page
Chapter 1 Introduction	1
Chapter 2 Efficient synthesis and expression of genes encoding protein triblock- copolymers	17
Chapter 3 Self-assembled and co-assembled structures of silk-collagen-like block copolymers	33
Chapter 4 Nucleation and growth by a silk-collagen-like block copolymer into supramolecular nano-ribbons	74
Chapter 5 Gelling kinetics and gel properties of silk-collagen-like block copolymers	91
Chapter 6 General discussion and conclusion	115
Summary	137
Samenvatting	141
Dankwoord	145
List of publications	149
Curriculum vitae	150
Educational activities	151

1

General introduction

A.A. Martens

1.1 Natural and man made polymers

Most animals shape their environment to some extent by building nests or burrows, but none do so to such extremes as humans. In the stone age, we used either stones or biologically derived materials like wood, bone and leather. We did not know it at the time but most bio-derived materials are crosslinked polymeric composite materials [1], with self-assembled micro and nano-structures, naturally tailored to optimally fulfill their purposes in the living beings that they once were part of. We took these natural materials, shaped them and used them in a way that seemed most fit to us.

Through time, humans improved their skills in shaping the world and materials around them through the cycle of primitive tools leading to better materials leading to better tools, and because of the simultaneous accumulation of knowledge. Today, we have surrounded ourselves with both natural and man made materials. We use them to shape the space around us to fit our human demands. The materials that we create are not only used in the constructions that we live in, and our daily objects and tools, but even in our bodies as e.g. surgical glue or implants. The common denominator is that all these products are intended to make our lives easier and/or more pleasant than living life in the wild, although this could be doubted some times.

Most of the materials serve their purpose pretty well, but the growing expectations that people have of their life standard, combined with a growing world population and limited resources demand new, efficient and advanced materials. In order for more people to have a better quality of life, new materials will have to be developed that perform equally at less cost or better than the materials they are replacing. The development of more efficient, new materials will lead to discoveries by observant researchers. In this way, new materials replacing old ones will lead to new applications which improve the quality of life.

Research in chemical synthesis and engineering have produced high performance synthetic polymers and composite materials, that have already started to replace metals like aluminum and steel. In volume (but not in weight) polymers have now outgrown the steel market. Almost half of the traded polymer volume is used for packaging. About 18% is used in the building industry and 5% is used in the automotive industry. Together, 15% is used for wires, cables, fibers, appliances and in household articles and the remaining 19% is used for various other, including specialty and high performance applications. Compared to the bulk

polymer market, the specialty polymer market is small in volume, but the price per product is much higher and the market volumes are growing with the expanding range of applications.

Demands on the structure and often multiple functions of new and efficient materials are now higher than ever. To improve on polymer properties and broaden applications even further, we would now like to make complex materials, tailored specifically to their purpose, with defined structures on every length scale. We are now able to study, understand and manipulate materials down to the atomic scale [2]. Still, manipulation and (mass)production of various functional materials with a defined nano- and meso-structure seems to be difficult, because we lack the proper tools. This brings us full circle, back to nature, that is able to produce the type of materials that we seek by self-assembly, but nature produces the materials for her own purposes. Therefore we have to harness biological processes to engineer the high complexity materials with the structures and functions that we desire. Also, biotechnological solutions may help in producing artificial biological materials more suited for medical applications.

1.2 Self-assembly

According to Wikipedia (which represents what we collectively think to be correct), “Self-assembly is a term used to describe processes in which a disordered system of pre-existing components forms an organized structure or pattern as a consequence of specific, local interactions among the components themselves, without external direction”. This can occur at all length scales [3], but the phenomenon is most interesting and the term is most commonly used for the organization of molecular units [4]. The reason is that self-assembly gives the tools to create extensive order on very small length scales (nm). This can either lead to organized nanoscopic objects built from smaller components, or to macroscopic materials with properties resulting from their nanostructure. Self-assembly on a macroscopic level is hardly used because we can rely on the many other tools that we have on this scale to shape our world.

For self-assembly to take place, both attraction and repulsion is needed within one molecule but they have to act on different parts of the molecule. There are different ways in which this can occur. For a molecule of two parts: A and B, there are three different cases: 1) A likes A, B likes B, and A repels B, 2) A likes A B repels B (or vice versa), and A repels B, 3) A repels A, B repels B, and A likes B, but not within one molecule because of steric

hindrance. The attraction will be responsible for the aggregation of the molecules while the repulsion will be responsible for the organization. An example could be a so-called amphiphilic molecule (**Figure 1.1 a**), with a head and a tail, for which the heads repel the tails, while the tails like each other very much [5]. If the head and tail would not be covalently linked, they would separate into two macroscopic phases. Because they cannot separate macroscopically, the heads and tails will both separate into their own nano-sized phase, while every head is still connected to its tail. Therefore, at least one dimension of such a phase is still on the same scale as the molecule itself [5]. The shapes that arise from such self-assembly vary enormously (**Figure 1.1**) (micelles, bilayers, lamellae, bicontinuous networks and others [5]), depending on the form of both head and tail, and on the amount of solvent used.

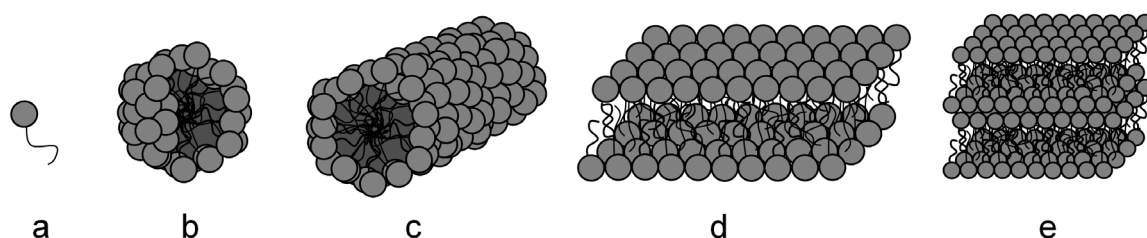


Figure 1.1 Examples of structures that may be expected from an (a) amphiphilic molecule with head repelling the tail: (b) micelle (c) worm like micelle (d) bilayer, (e) lamellar phase.

The typical feature of self-assembly is that it happens by itself. Therefore the formed structures reside in an energy minimum. A free molecule in solution has a higher Gibbs free energy than the same molecule in the structure would have. Therefore the structure is formed spontaneously. It might be possible however that there are several (metastable) energy minima, all associated with their own structure but that the process along which the structure is formed dictates in which minimum the structure arrives [6]. Under certain conditions the structures may be kinetically trapped (or frozen), and under other conditions the structures may be interconvertible. Usually, small molecules are dynamic and form one structure at a time. Therefore, structures of small molecules may easily be converted into each other by shifting physical circumstances like temperature, pH, and concentrations. The occurrence of different metastable structures at the same time is more likely for large molecules (polymers). The larger the polymer, the more complex the self-assembly is. This means that there can be more possible structures with similar minimal energies and that the molecules can get kinetically trapped in several of these minima under the same circumstances [6].

Often the word self-assembly is reserved for the spontaneous assembly of equal units into an organized structure. The word “self” in this case does not only refer to the assembly happening by itself, but also to the spontaneous assembly happening between equal units i.e. “They assemble with their own kind”. Spontaneous assembly of different components, as is the case for different subunits of a protein complex is then dubbed “co-assembly”.

Biology is littered with examples of self- and co-assemblies, from nanoscopic functional objects like co-assembled enzyme complexes, often present in or on the co-assembled phospholipid membranes that compartmentalize cells, to macroscopic nanostructured materials, like plant cell wall [1]. Some of the structures, like oyster shell and bone also include minerals [1]. However, in a chemical classification, there are only two major polymer groups taking part in these structures: proteins and polysaccharides (which can also fulfill other than structural functions). Other biopolymers include the nucleic acids DNA and RNA which mainly have a biological data storage and transfer function, which also are interesting for the structures that they can form [7, 8], and some other polymers for energy storage (polyesters [9] and branched non-ribosomal polypeptides [10]), that may be useful as bio-derived bulk polymers.

Compared to biology, humans have only recently started to produce self-assembling molecules (synthetically). The smallest and simplest ones are surfactants [5]. They assemble in water to form micelles and vesicles because of the interactions of their hydrophobic tails. The hydrophilic and sometimes charged surfaces of such objects prevent them from aggregating even further. Still, small surfactants in a larger structure are not fixed in place. They can dynamically enter and leave the structure [11]. Therefore the structures formed by surfactants should maybe be called self-organized and not yet self-assembled. At higher concentrations surfactants can form many different nano-structures [5]. Self-organization becomes self-assembly when aggregating (self-assembling) blocks stay in place. This can be achieved with large aggregating polymer blocks, and strong interactions, like hydrogen bonding or Coulombic interactions.

Larger than surfactants are the block copolymers: different, phase separating polymers grafted to each other to form one molecule. The most simple ones are diblock copolymers, molecules of only two different polymer blocks. They tend to behave similarly to, and could be regarded as surfactants, forming dynamic micelles and various nanostructures [5]. Actually

a classical surfactant could be regarded as a diblock copolymer, of which the hydrophilic block is extremely small.

Triblock copolymers of two different blocks can be separated into two groups. The first has one aggregating block in the middle (ABA) and the second contains two aggregating blocks on the flanks (BAB). An ABA triblock copolymer is also likely to self-organize into micelles, vesicles, lamellae, bicontinuous, and various other phases [12]. The BAB block copolymers can form similar structures which characteristically also have loops of hydrophilic middle blocks when the hydrophobic end blocks are located in the same phase [13]. A cartoon of a micelle with a surface of such hydrophilic loops (**Figure 1.2 a**) would be reminiscent of a children's drawing of a flower [13]. At increasing concentrations they form an additional structure in which the two sticky blocks of one molecule can sit in different micelle cores, forming networks of these flower-like micelles (**Figure 1.2 b**) [13]. As for any surfactant, variation of block sizes of any of the above block copolymers leads to variation in the structures formed. Still most block copolymer structures have a dynamic character, although this decreases with increasing block size of the aggregating block [14].

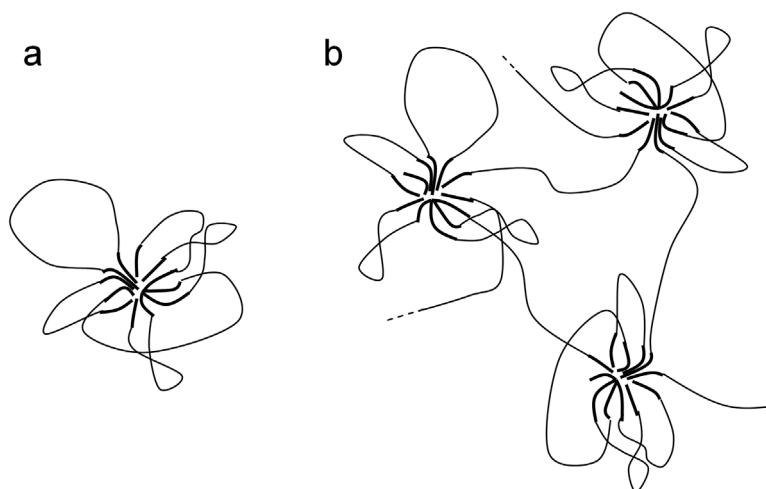


Figure 1.2 (a) Cartoon of a flower like micelle from a triblock copolymer with a corona of loops and "sticky" end blocks depicted as thicker lines in the micellar core. (b) Cartoon of a network of such flower like micelles, where endblocks of one molecule reside in different micellar cores.

Monodispersity is thought to benefit self-assembly because equally sized building blocks fit perfectly to each other, reducing the number of defects in the larger self-assembled structure. To build simple self-assembled structures one could use block copolymers with monodisperse homopolymer blocks. If, however, the blocks were to be made of various

monomers, over which we had the control over the exact monomer sequence, an array of different shapes and structures would be accessible. This would lead to several levels of self-assembly. The first one would be the folding of the polymer chain itself, into blocks that have their own shape and conformation. Then, the blocks of different molecules could interact with each other forming larger supramolecular structures. The structure and properties of such a synthetic block copolymer would come quite close to those of natural proteins. However, with classical polymerization techniques monodisperse blocks are already difficult to obtain, and long sequential polymers are unobtainable. Molecules with different sequential blocks have been created before, by grafting synthetic polypeptides to homopolymers [15], but the low amounts obtained are only useful for fundamental research and not for material science. A more successful approach is the grafting of polymers to biotechnologically produced polypeptides [16], which gives better production yields. Possibly the best way of obtaining large amounts of such molecules would be to biotechnologically produce an entire block copolymer as one large protein, comprising different polypeptide blocks with different physical properties.

1.3 Proteins

The prime example of a natural monodisperse sequential polymer is a protein. A primary gene product contains up to 20 different amino acids (monomers) in the primary structure (defined sequence). All protein molecules encoded in the same gene (template) are identical, with exactly the same sequence and length. The polymerization of natural amino acids into a protein creates a polyamide backbone that is equivalent to nylon 2 (**Figure 2.3**). Depending on which amino acids are used in the primary sequence, the polyamide has side chains side that define the polymer properties. The amino acid side chains can vary in size and polarity. Some contain a positive or a negative charge. Some have hydrogen bond donors or acceptors. They can also be aromatic, or aliphatic, both branched and unbranched. Some can be chemically modified after polymerization, either in the cell or by organic chemistry. These posttranslational modifications create a few additional amino acid residues. All these different properties in a sequence result in a polymer (protein) with preferential ways of folding. Therefore, under the same circumstances, all protein molecules of the same type, have the same conformation (3D structure) in equilibrium with a unique biological function. As a

result of various physical or chemical stimuli, this conformation may change uniformly for all identical protein molecules.

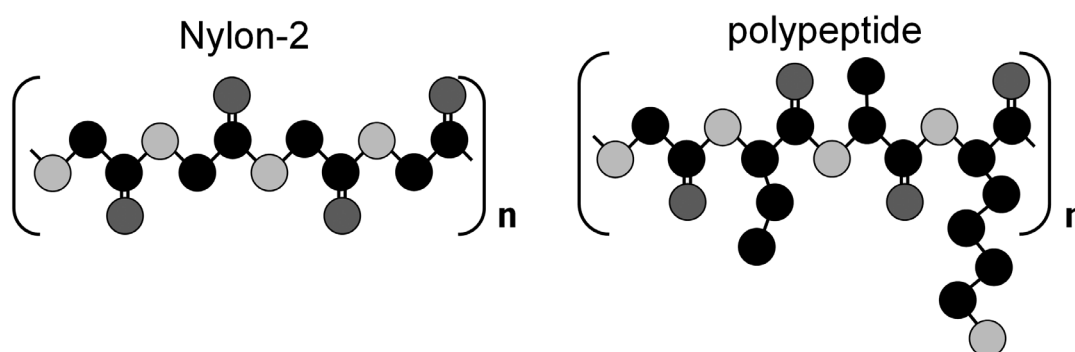


Figure 2.3 A cartoon of the structure of Nylon-2 compared to a polypeptide. Carbon atoms are black, oxygen dark gray and nitrogen light gray. Both are polyamides. Nylon-2, lacking side groups, is exactly the same as the backbone of a polypeptide or protein, and could be regarded as polyglycine.

Every natural protein has its unique functions (either inside or outside the organism). There are many different primary structures with accompanying conformations, that fulfill a multitude of biological functions like catalysis, transport of various substances, signal transduction, and actuation. Since we are interested in material properties, we would like to emphasize that several proteins are structural elements. One example is actin, the major constituent of the cytoskeleton which has many functions inside the living cell, and dominates its elastic properties [17-19]. An other example is collagen [20], a major component of the extracellular matrix, giving elastic properties to tissues, tendons, bone cartilage and skin. Elastin [21] has a similar role, but is mainly found in connective tissue and skin. It allows these tissues to take their old form after deformation. A different, example is fibrin [22], which usually is soluble, but forms fibers when blood clotting is triggered. Spider silk [23] *Bombyx mori* silk [24] and mussel byssal threads [25] are typical examples of structural proteins that fulfill their purpose as super strong fibers outside the organism.

Especially the silks and mussel byssal proteins, consists of several different blocks with different structural functions [25]. Some blocks interact with other blocks within the same molecule, promoting a protein to fold into its conformation. Some assemble with blocks from other, identical molecules to form a larger self-assembled structures, and some interact specifically with different molecules to form co-assemblies. Such are the polymers that intend to emulate. We aim to build highly defined, large, self-assembled, supramolecular structures, starting from single molecules of one molecular type.

For all the protein molecules of a single type to be identical, there must be a template. This is the DNA, which serves as a large databank for all the proteins that the cell might ever need in its lifetime. Usually the DNA contains only a single copy (gene) encoding a certain protein. The so-called promoter region of the gene, and the stability of the m-RNA contains information on how much of this protein should be synthesized under which conditions. The natural protein production machinery works as follows. When a certain protein is needed, transcription factors bind to or unblock the promoter region of the corresponding gene and recruit RNA polymerase to transcribe the gene into many copies of RNA. These serve to transport the product information from the DNA (data storage) to the ribosome (protein factory). Every copy of RNA is read several times by the ribosomes and with every reading one protein molecule is produced. In short: one gene leads to many copies of RNA, each of which leads to many more copies of identical protein molecules.

For two reasons, protein production is more versatile and controlled than chemical polymerization reactions, be it natural or entirely synthetic. The use of a template (DNA) encoding a polyamide enables production of identical sequential polymers in only one production step (one fermentation).

1.4 Protein polymers

The natural protein production machinery can be used to produce monodisperse polymers with an identical sequence in a highly controlled fashion [26-28]. These molecules include transgenically produced natural and modified proteins, but also designed block copolymers can be produced in this way [26, 27]. The blocks themselves may be repetitive sequences of different amino acids.

This biotechnological approach (in some variety), for producing sequential and monodisperse protein polymers, was pioneered in the beginning of the 90ies by J. Cappello and F. Ferrari [26] and the group of D.A. Tirrell [28]. The term “protein polymer” was first used by Cappello [26]. Their first molecules were diblock copolymers [26] containing silk-like and elastin-like amino acid sequences. Since then, many different designed nature inspired protein (block-co)polymers have been produced [27], comprising many different sequential polymers, often combined to form di- and multiblock copolymers.

Tirrells group produced the first silk-like protein polymer carrying only negative charge [29]: a repeating octapeptide (GAGAGAGE)₃₆ with small flanking sequences. This molecule

was inspired on the naturally occurring (glycine-alanine)_n repeats that are responsible for crystalline physical crosslinks in natural silk [30-32]. When crystallized from a mixture of methanol and formic acid, this sequence produces needle shaped lamellar crystals [29], in which the molecules form stacks of β -sheets. However when exposed to water (or water vapor), the conformation of the molecules changes into a structure which has not yet been fully resolved [33, 34] [35]. Still, in water, this molecule should be able to switch between being soluble and non soluble depending on pH or opposite charges. Therefore we would like to use this repetitive amino acid sequence as a block type in novel pH and (poly)electrolyte responsive, self-assembling, and co-assembling block copolymers.

Protein polymers are expensive relative to simple chemosynthetic (homo)polymers. Therefore the applications are, for now, limited to high value applications like medical applications, and to thin film applications in which a small amount of protein polymer is sufficient to cover a large surface. Most of the protein polymers that have been produced to date are intended to fulfill a medical application [36, 37], for example, injectable gels for controlled release or tissue engineering scaffolds. They often contain two different block types: a structural block, responsible for self-assembly and structure formation, and a block that interacts with the living cells. The block responsible for self-assembly often leads to gel formation under physiological conditions.

Control over the gelling behavior and rheology of these polymers would be very useful since injecting or casting requires low viscosity, after which a gel may develop to retain its form. However, in aqueous environment, most of the protein (block-co)polymers spontaneously form gels, except for the (GAGAGAGE) repeat [29] that is negatively charged and therefore self-repelling and hydrophilic. This molecule does not form aggregates or gels in aqueous solution at physiological pH. High solubility is desired during material processing, after which the material is allowed to set.

1.5 The FITAPEP project

This Dutch Polymer Institute (DPI) funded project, aims to lay the basis for the development of super strong fibers with improved transversal strength with respect to already existing ones. Such fibers may be realized with bio-inspired block copolymers, containing monodisperse, sequential blocks (e.g. charged ones) that self-assemble upon a stimulus (e.g. a change in pH), to form a self-assembled gel. While expelling the solvent from the gel, such a gel may be

spun in order to align the polymers into a strong fiber, in which the self-assembled blocks are responsible for the improved transversal strength while other more flexible blocks bear the load in the fiber direction.

A similar approach may be to use 2 oppositely charged block copolymers in which the self-assembling blocks carry the charge. Both components would be soluble at neutral pH, but when mixed, they would co-assemble into a gel that may be spun into a fiber with improved transversal strength.

To investigate these approaches, several steps have to be taken. A method has to be developed to quickly produce a large amount (>500 mg for research, more for applications) of monodisperse, sequential block copolymers with charge containing self- and co-assembling blocks. The gels resulting from self- and co-assembly of the various products have to be studied with respect to their structure and structure formation kinetics, to assess their suitability for gel spinning. Candidate gels for gel spinning will have to be selected and material processing optimized.

During this exploratory research, we will come across various self- and co-assembled gels with sometimes surprising nano-structures and material properties that may have more direct medical and technical applications.

1.6 Aim of this thesis

The aim of this thesis is to study the properties of several large, water-soluble (monodisperse) protein triblock copolymers, with various sequential, either positively or negatively charged blocks, that self-assemble in response to a change in pH, or co-assemble with oppositely charged polyelectrolytes (including each other). The study will focus on the effect that such molecules have on the kinetics of structure formation, the morphology of the self-assembled and co-assembled structures, and on their associated (gel)material properties. To obtain a variety of such molecules, in sufficient amounts for material testing, a modular cloning system will be developed to clone various blocks and genes encoding the various block copolymers that will be expressed in the yeast *Pichia pastoris* [38] because it offers a promising avenue for producing large amounts of repetitive protein polymers. This is done in the light of developing molecules that are likely to form super strong fibers with improved transversal strength after material processing.

1.7 Outline of this thesis

The pH and charge responsive, self-assembling (GAGAGAGE)_n silk-like sequence [29] was combined with a monodisperse, biodegradable, biocompatible and hydrophilic collagen-like sequence, previously called P4 [39], to form block copolymers. Although we are used to collagen-like sequences forming gels, P4 is highly hydrophilic and remains soluble under most conditions (various pH and T). The resulting molecules were 802 amino acid long block copolymers with self-assembling and non-assembling blocks.

To build oppositely charged block copolymers, we chose to use the same block combination of the collagen-like sequence [39] with a new, positively charged silk-like sequence, almost identical to the one mentioned above. In the new silk-like sequence, the positively charged histidine replaced the negatively charged glutamic acid. Both the negatively charged silk-like block and the positively charged silk-like block could be combined on a genetic level with the collagen-like block, to encode various complementary, and oppositely charged block copolymers. In **Chapter 2** we describe the combination of DNA template blocks encoding the silk-like sequences and the collagen-like sequence to form genes that encode the different block copolymers. We also describe the block copolymer production by fermentation and their purification.

Once pure, the various nano- and meso-structures that they form upon self- and co-assembly could be studied. In **Chapter 3** we investigate the various structures formed on the nano to meso scale using SAXS, DLS and various microscopic techniques like (cryo-)TEM and AFM. The experimental data were compared to MD models generated by a group with which we currently cooperate.

In **Chapter 4**, we studied the rate at which some of the self-assembled structures are formed using time resolved CD measurements at 200 nm. Formation kinetics may be relevant to future material processing. The self-assembled structures formed macroscopic gels. Some gel properties, also described in **Chapter 4**, like the absence of swelling in water and the pH at which the gel melts (determined with DLS) together with microscopy (AFM and TEM) contributed to the understanding of the physical nature of these self-assembled structures.

In **Chapter 5** we tested the mechanical properties of self-assembled gels using mechanical spectroscopy in a rheometer. These mechanical properties varied as a function of time after sample preparation, polymer concentration, pH, and the type of polymer product. The products differed in block order and/or block charge, with either glutamic acid or

histidine in the silk-like block. Also the properties as a function of temperature were tested and compared to CD measurements. We also investigated, as a function of temperature, structural changes responsible for changes in material properties. This was done by comparing temperature dependent CD measurements to temperature dependent mechanical spectroscopy. We linked the macroscopic structural properties to the self-assembled nanostructures.

The review discussion of this thesis in **Chapter 6**, explains the molecular design considerations of our block copolymers in detail with respect to the original project aim (super strong fibers) and the consequences of the design for the choice of production method (**Chapter 2**), the structures formed by the molecules (**Chapter 3**), the kinetics with which some of these structures form (**Chapter 4**), and the material properties of the gels that these structures constitute (**Chapter 5**). In **Chapter 6** the best candidate-gel for gel spinning is pointed out, more direct applications are suggested for the tested gels, and other directions of research involving structure formation by our self- and co-assembling molecules are considered.

References

- 1 Meyers, M. A., Chen, P. Y., Lin, A. Y. M. and Seki, Y. (2008) Biological materials: Structure and mechanical properties. *Progress in Materials Science* **53**, 1-206
- 2 Lyo, I. W. and Avouris, P. (1991) Field-Induced Nanometer-Scale to Atomic-Scale Manipulation of Silicon Surfaces with the Stm. *Science* **253**, 173-176
- 3 Whitesides, G. M. and Grzybowski, B. (2002) Self-assembly at all scales. *Science* **295**, 2418-2421
- 4 Park, C., Yoon, J. and Thomas, E. L. (2003) Enabling nanotechnology with self assembled block copolymer patterns. *Polymer* **44**, 6725-6760
- 5 Soler-illia, G. J. D., Sanchez, C., Lebeau, B. and Patarin, J. (2002) Chemical strategies to design textured materials: From microporous and mesoporous oxides to nanonetworks and hierarchical structures. *Chemical Reviews* **102**, 4093-4138
- 6 Keller, A. and Cheng, S. Z. D. (1998) The role of metastability in polymer phase transitions. *Polymer* **39**, 4461-4487
- 7 He, Y., Ye, T., Su, M., Zhang, C., Ribbe, A. E., Jiang, W. and Mao, C. D. (2008) Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198-U141
- 8 Rothmund, P. W. K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302
- 9 Lee, S. Y. (1996) Bacterial polyhydroxyalkanoates. *Biotechnology and Bioengineering* **49**, 1-14
- 10 Simon, R. D. (1971) Cyanophycin Granules from Blue-Green Alga *Anabaena-Cylindrica* - Reserve Material Consisting of Copolymers of Aspartic Acid and Arginine. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 265-&
- 11 Hayter, J. B. and Penfold, J. (1981) Self-Consistent Structural and Dynamic Study of Concentrated Micelle Solutions. *Journal of the Chemical Society-Faraday Transactions I* **77**, 1851-1863

- 12 Alexandridis, P., Olsson, U. and Lindman, B. (1998) A record nine different phases (four cubic, two hexagonal, and one lamellar lyotropic liquid crystalline and two micellar solutions) in a ternary isothermal system of an amphiphilic block copolymer and selective solvents (water and oil). *Langmuir* **14**, 2627-2638
- 13 Xu, B., Yekta, A., Li, L., Masoumi, Z. and Winnik, M. A. (1996) The functionality of associative polymer networks: The association behavior of hydrophobically modified urethane-ethoxylate (HEUR) associative polymers in aqueous solution. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **112**, 239-250
- 14 Semenov, A. N., Joanny, J. F. and Khokhlov, A. R. (1995) Associating Polymers - Equilibrium and Linear Viscoelasticity. *Macromolecules* **28**, 1066-1075
- 15 Cho, I., Kim, J. B. and Jung, H. J. (2003) Synthesis and characterization of di- and triblock copolymers of poly(ethylene oxide) and poly(DL-valine-co-DL-leucine). *Polymer* **44**, 5497-5500
- 16 Smeenk, J. M., Otten, M. B. J., Thies, J., Tirrell, D. A., Stunnenberg, H. G. and van Hest, J. C. M. (2005) Controlled assembly of macromolecular beta-sheet fibrils. *Angewandte Chemie-International Edition* **44**, 1968-1971
- 17 Gardel, M. L., Shin, J. H., MacKintosh, F. C., Mahadevan, L., Matsudaira, P. and Weitz, D. A. (2004) Elastic Behavior of cross-linked and bundled actin networks. *Science* **304**, 1301-1305
- 18 Janmey, P. A., Hvidt, S., Lamb, J. and Stossel, T. P. (1990) Resemblance of Actin-Binding Protein Actin Gels to Covalently Cross-Linked Networks. *Nature* **345**, 89-92
- 19 Storm, C., Pastore, J. J., MacKintosh, F. C., Lubensky, T. C. and Janmey, P. A. (2005) Nonlinear elasticity in biological gels. *Nature* **435**, 191-194
- 20 Vanderrest, M. and Garrone, R. (1991) Collagen Family of Proteins. *Faseb Journal* **5**, 2814-2823
- 21 Debelle, L. and Tamburro, A. M. (1999) Elastin: molecular description and function. *International Journal of Biochemistry & Cell Biology* **31**, 261-272
- 22 Doolittle, R. F. (1984) Fibrinogen and Fibrin. *Annual Review of Biochemistry* **53**, 195-229
- 23 Gosline, J. M., Demont, M. E. and Denny, M. W. (1986) The Structure and Properties of Spider Silk. *Endeavour* **10**, 37-43
- 24 Lotz, B. and Cesari, F. C. (1979) Chemical-Structure and the Crystalline-Structures of Bombyx-Mori Silk Fibroin. *Biochimie* **61**, 205-214
- 25 Coyne, K. J., Qin, X. X. and Waite, J. H. (1997) Extensible collagen in mussel byssus: A natural block copolymer. *Science* **277**, 1830-1832
- 26 Cappello, J., Crissman, J., Dorman, M., Mikolajczak, M., Textor, G., Marquet, M. and Ferrari, F. (1990) Genetic-Engineering of Structural Protein Polymers. *Biotechnology Progress* **6**, 198-202
- 27 Haider, M., Megeed, Z. and Ghandehari, H. (2004) Genetically engineered polymers: status and prospects for controlled release. *Journal of Controlled Release* **95**, 1-26
- 28 McGrath, K. P., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1992) Genetically Directed Syntheses of New Polymeric Materials - Expression of Artificial Genes Encoding Proteins with Repeating (Alagly)₃proglugly Elements. *Journal of the American Chemical Society* **114**, 727-733
- 29 Krejchi, M. T., Atkins, E. D. T., Waddon, A. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1994) Chemical Sequence Control Of Beta-Sheet Assembly In Macromolecular Crystals Of Periodic Polypeptides. *Science* **265**, 1427-1432
- 30 Bini, E., Knight, D. P. and Kaplan, D. L. (2004) Mapping domain structures in silks from insects and spiders related to protein assembly. *Journal of Molecular Biology* **335**, 27-40
- 31 Lotz, B. and Keith, H. D. (1971) Crystal Structure of Poly(L-Ala-Gly)Ii - Model for Silk I. *Journal of Molecular Biology* **61**, 201-202
- 32 van Hest, J. C. M. and Tirrell, D. A. (2001) Protein-based materials, toward a new level of structural control. *Chemical Communications*, 1897-1904
- 33 Chen, C. C., Krejchi, M. T., Tirrell, D. A. and Hsu, S. L. (1995) Effect of Water on the Structure of a Model Polypeptide. *Macromolecules* **28**, 1464-1469

- 34 Topilina, N. I., Higashiya, S., Rana, N., Ermolenkov, V. V., Kossow, C., Carlsen, A., Ngo, S. C., Wells, C. C., Eisenbraun, E. T., Dunn, K. A., Lednev, I. K., Geer, R. E., Kaloyeros, A. E. and Welch, J. T. (2006) Bilayer fibril formation by genetically engineered polypeptides: Preparation and characterization. *Biomacromolecules* **7**, 1104-1111
- 35 Werten, M. W. T., Moers, A. P. H. A., Vong, T., Zuilhof, H., van Hest, J. C. M. and de Wolf, F. A. (2008) Biosynthesis of an Amphiphilic Silk-Like Polymer. *Biomacromolecules*
- 36 Deming, T. J. (1997) Polypeptide materials: New synthetic methods and applications. *Advanced Materials* **9**, 299-&
- 37 Megeed, Z., Cappello, J. and Ghandehari, H. (2002) Genetically engineered silk-elastinlike protein polymers for controlled drug delivery. *Advanced Drug Delivery Reviews* **54**, 1075-1091
- 38 Werten, M. W. T., Van den Bosch, T. J., Wind, R. D., Mooibroek, H. and De Wolf, F. A. (1999) High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**, 1087-1096
- 39 Werten, M. W. T., Wisselink, W. H., van den Bosch, T. J. J., de Bruin, E. C. and de Wolf, F. A. (2001) Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Engineering* **14**, 447-454

2

Efficient synthesis and expression of genes encoding protein triblock copolymers

A.A. Martens, M.W.T. Werten, G. Eggink, M.A. Cohen Stuart, F.A. de Wolf

Abstract

Block copolymers, containing blocks with different physical properties have found high value applications like nano-patterning and drug delivery. By gaining control over the exact monomer sequence and length, applications could be expanded. However, large, sequential polymers are not obtainable with classical organic chemistry. Therefore we constructed synthetic genes, encoding designed, long amino acid sequences and employed the natural protein production machinery of the yeast *Pichia pastoris* to produce them. We describe the secreted production in yeast of the 65.7 kDa 'CS^ES^EC' and 'S^ECCS^E' and of the 66.1 kDa 'CS^HS^HC' and 'S^HCCS^H'. These four products are all triblock copolymers, consisting of: 1) silk-like (S) blocks that self-assemble depending on the pH, and 2) biocompatible collagen-like (C) blocks that do not self-organize. The (silk-like) S blocks consist of an octapeptide repeat sequence ((GA)₃GX)_n in which X is either glutamic acid ("S^E") or histidine ("S^H") Both are soluble at neutral pH, while S^E and S^H self-assemble at low and high pH respectively. The product yields are in the gram per liter range, such that various applications of these promising biomaterials become possible. pH-Responsive self-assembly of the S blocks of all four polymers results in the formation of transparent gels, for the S^E containing products even at a concentration as low as 0.9 g/l.

2.1 Introduction

Stimulus-responsive, nano-structured, and self-assembling polymer materials are essential for high value applications like self-healing coatings, chemo-mechanical fibers, sensors, smart packaging, surgery, regenerative medicine and pharmaceuticals [1-7]. Speed and precision of stimulus-induced supramolecular self-assembly are expected to benefit from the use of polymers built from chiral monomers and consisting of one single molecular type with exactly defined length, domain structure and monomer sequence. While chemically synthesized polymers normally lack these features, they are the hallmark of biosynthetic proteins. In addition, chemical polymers typically lack protein-borne biocompatible features like cell attachment sites, or programmed biodegradation, exploitable in regenerative medicine and pharmaceuticals. These considerations are the basis of the rapidly expanding field of protein polymer science, focusing on nature-inspired designer proteins with a structure-forming function [8].

In the present work we focus on the production and purification of four novel, entirely biosynthetic triblock copolymers consisting of silk-like blocks and collagen-like blocks. The DNA template of the different blocks were combined to form genes, coding for monodisperse, sequential protein block copolymers. The genes were expressed by the yeast *Pichia pastoris* and the excreted products were purified from the fermentation broth by selective salt and solvent precipitations respectively, with varying yields, all in the range of a gram per liter of broth.

For our design, we selected two basic block types. One block (S, for silk-like) is a 192 amino acid long pH-responsive silk-like octapeptide repeat of glycine (G) and alanine (A) [9], (GAGAGAGX)₂₄, with X being either glutamic acid (E) or histidine (H). The repeat in the S block is capable of forming crystalline stacks of antiparallel AGAGA β -sheets [9], bordered by GXG γ -turns [9, 10]. At neutral pH, i.e. in the charged state, the S chain assumes a random - extended conformation and is well-soluble in water. Conversely, in their uncharged state, at high pH for the block containing histidine (S^H), and at low pH for the block containing glutamic acid (S^E), they are insoluble. The other block (C for collagen-like) is a 198 amino acid long extremely hydrophilic glutamine-, asparagine- and serine-rich collagen-like designer polypeptide that our group developed [11]. It has a strong preference for unordered structure at all pH and does not form supramolecular collagen-like assemblies, due to an

absence of prolyl hydroxylation [11]. It has the ability to direct human cells in culture selectively to patches of a substrate that are coated with the polypeptide [12] and has favorable biocompatibility as compared to animal collagen-derived products in blood transfusion applications [13].

Four polymers were designed to self- and co-assemble under various conditions. Two complementary arrangements of the blocks were produced, namely two consecutive silk-like blocks in the middle flanked by two collagen-like blocks at the N- and C-terminal ends of the molecule (CSSC) and vice versa (SCCS). Of both arrangements, we produced both a negatively charged version (containing S^E) and a positively charged version (containing S^H), amounting to four products with complementary arrangements and complementary charge. Self-assembly to form gels and materials is expected for the S^E containing products at low pH and for the S^H containing products at high pH. Co-assembly at neutral pH may occur when S^E and S^H containing products are mixed.

For rapid building of genes encoding such block copolymers, PCR could not be used, because of the repetitive nature of these genes. For such a gene, hybridization of single stranded DNA would be possible in many ways, and consequently lead to faulty products. Instead, to create the different S blocks, we annealed complementary oligonucleotides (76 base pairs) and, in a plasmid, connected them to each other through restriction and ligation. The genes encoding the whole block copolymers were created similarly by restriction and ligation of the newly synthesized S blocks and the C block.

The restriction enzymes that we used (BsaI and BanI), allowed us to efficiently and seamlessly enlarge the S block and to connect the S and the C blocks to each other. BsaI cuts next to its recognition site, cutting in a DNA sequence in the block that codes for glycine and alanine. If oriented properly, enlarging the block by restriction and ligation results in the BsaI recognition site not being included in, but only at the edge of the newly formed block. This enables quick enlargement of the block, multiplying the block size by a factor of two in every cloning cycle. The middle two nucleotides of the BanI recognition- and cutting-site can be chosen so the DNA on the end of the block also codes for the amino acids glycine and alanine. This can make the edges of several silk- and collagen-like blocks fit seamlessly.

Finally, all four triblock copolymers were produced at high yield in the methylotrophic yeast *Pichia pastoris* similarly to previous recombinant gene expression [13, 14]. We chose this eukaryote because repetitive DNA can suffer from recombination in prokaryotes like

E.coli [15] and because *P.pastoris* has a good track record of producing repetitive amino acid sequences [11, 13, 16, 17]. Afterwards, purification using selective precipitation with solvent and salt, was an easy and scalable procedure.

2.2 Materials and methods

2.2.1 Generation of DNA template blocks, genes and recombinant strains

The first template block, encoding the S^E block silk-like sequence (GAGAGAGE)₂₄, was produced as follows. A double-stranded adapter was constructed by annealing of oligonucleotides:

5'AATTCGGTCTCGGTGCTGGTGCTGGTGCTGGTGCTGGTGAGGGAGCCGGTGCTGGAGCCG GCGAAGGTGCCTAAGCGGCCGC3' and 5'TCGAGCGGCCGCTTAGGCACCTT-CGCCGGCTCCAGCACCGGCTCCCTCACCAGCACCGAGACCG3'.

The adapter was then ligated into an *EcoRI/XhoI* digested, modified pMTL23 vector [18] called pMTL23-Δ*BsaI*, from which the normally present *BsaI* site had been removed. The insert was elongated to encode 24 repeats of the amino acid sequence (GAGAGAGE) by digestion with *BsaI/BanI* and directional ligation. Proper length of the block while elongating was verified with colony PCR.

The second template block, encoding the S^H block silk-like sequence (GAGAGAGH)₂₄, was produced exactly as described for the S^E block but with the oligonucleotides: 5'AATTCGGTCTCGGTGCTGGTGCTGGTGCTGGTGCTGGTGACGGAGCCGGTGCTGGAGCCG GCCATGGTGCCTAAGCGGCCGC3' and 5'TCGAGCGGCCGCTTAGGCACCATG-GCCGGCTCCAGCACCGGCTCCGTGACCAGCACCGAGACCG3', which encode for histidine on the X position of the (GAGAGAGX) repeat instead of glutamic acid.

The third template block, encoding the C-block hydrophilic collagen-like sequence was produced as follows. A double-stranded adapter was constructed by annealing of oligonucleotides:

5'AATTCGGTCTCGGTGCTGGTGCAACCCGGTGAGGGTGCCTAAGCGGCCGC3' and 5'TCGAGCGGCCGCTTAGGCACCCTCACCGGGTGCACCAGCACCGAGACCG3'. The adapter was then inserted into the *EcoRI/XhoI* sites of the pMTL23-Δ*BsaI* vector described above. The resulting vector was linearized with *DraIII* and dephosphorylated. The gene encoding the hydrophilic collagen-like sequence (P2) was cut from the previously described

vector pMTL23-P2 [11] with *DraIII/Van9II* and inserted into the linearized vector, creating the C-block template.

Before they were combined to form genes, all DNA blocks were sequenced to verify identity, correct frame and intactness of the block. *BsaI* and *BanI* were used for digestion and recursive directional ligation of the template blocks, first into diblocks CS^E , S^EC , CS^H and S^HC and then into tetrablocks CS^ES^EC , S^ECCS^E , CS^HS^HC and S^HCCS^H . Finally, each of the four tetrablocks were cloned into expression vector pPIC9 (Invitrogen) using *EcoRI* and *NotI*. The resulting vectors were linearized with *SaII* to promote homologous integration at the *his4* locus upon transformation of *Pichia pastoris* GS115 by electroporation, as described previously [13]. When template blocks were designed, codons were used randomly, except for the ones not preferred by *P.pastoris* [19], and methylation sites GATC and CCWGG were avoided at restriction endonuclease sites.

2.2.2 Protein polymer production and purification

Polymer production was obtained in fed-batch fermentations of *Pichia pastoris* in 2.5-liter Bioflo 3000 bioreactors (New Brunswick Scientific), essentially as previously described [11]. Throughout the fermentations, the pH was maintained at 5 for S^E containing products and the pH was maintained at 3 for S^H containing products. The methanol concentration was maintained at 0.2 % (v/v) during the induction phase. The polymers were secreted into the fermentation medium, which was separated from the cells by 15 minutes centrifugation at 20,000 g and 4°C (in a Sorvall centrifuge with a SLA1500 rotor), followed by microfiltration of the supernatant.

The glutamic acid bearing S^E containing polymers were precipitated selectively from the fermentation supernatant by adding ammonium sulphate to 30 % saturation (at 4°C), incubating for 30 min at 4 °C and centrifugation for 20 min. at 8000 g and 4°C (Sorval, SLA1500). The polymer pellets were dissolved in 20% of the original volume of 10 mM ammonia (pH 9) and the precipitation procedure was repeated once. The resuspended polymers were selectively precipitated by adding acetone to a final concentration of 80% (v/v). Resuspension and acetone precipitation were repeated once more, after which the pellets were resuspended in water and freeze-dried for storage. The salt containing freeze-dried products were each resuspended in 100 ml 50 mM ammonia and extensively dialyzed against 10 mM ammonia, after which the desalted products were freeze-dried again.

The histidine bearing S^H containing polymers were precipitated selectively from the fermentation supernatant by adding ammonium sulphate to 45 % saturation (at 4°C), incubating for 30 min at 21 °C and centrifugation for 20 min. at 8000g and 4°C (Sorval, SLA1500). The polymer pellets were dissolved in 20% of the original volume of 100 mM acetic acid from which the polymers were selectively precipitated by adding acetone to a final concentration of 50% (v/v). The pellets were dissolved in 300 ml 10 mM acetic acid and freeze-dried for storage. The salt containing freeze-dried products were resuspended in 100 ml 50 mM formic acid and extensively dialyzed against 10 mM formic acid, after which the polymers were freeze-dried again.

2.2.3 Product identification and purity assessment

During the purification procedure, the purity on the protein level was assessed on SDS-PAGE (Invitrogen NuPAGE Novex). Conductivity measurements were used to assess the amounts of salt in the purified products by comparing the conductivity of dissolved products with the conductivity of the solvent and attributing the difference to $(NH_4)_2SO_4$. Amino acid content analysis after protein hydrolysis was used to assess purity of the final product (performed by Ansynth service b.v. the Netherlands). The (poly)saccharide content was determined by a phenol-sulfuric acid sugar assay [20].

2.2.4 Gel formation induced by shifting the pH

Solutions of all products were made by dissolving $CS^E S^E C$ and $S^E CCS^E$ at 0.9 g/l in 1 mM NaOH, and $CS^H S^H C$ and $S^H CCS^H$ at 0.9 g/l in 1 mM HCl. Of these solutions, 4 ml were transferred to a separate vial and to induce aggregation, 40 μ l of 1M HCl was added to the $CS^E S^E C$ and $S^E CCS^E$ solutions and 40 μ l of 1M NaOH was added to the $CS^H S^H C$ and $S^H CCS^H$ solutions. For $CS^H S^H C$ and $S^H CCS^H$ a second preparation was done containing 4 ml 4.5 g/l solution and 40 μ l 1M NaOH. The solutions were allowed to gel for 24 h before turning the vials upside down (**Fig. 2.5**). Photos were taken half an hour after turning.

The coding part of the annealed, double stranded, DNA sequence is flanked by the endonuclease recognition sites of *BsaI* and *BanI*, (**Fig. 2.1 a**) which were used to enlarge blocks to the desired size and to connect different blocks. *BsaI* is an unusual restriction endonuclease because it does not cut the DNA strands at, but next to the recognition site (**Fig. 2.1 a**). This has two advantages for enlarging repetitive blocks. The first is that *BsaI* cuts the DNA sequence unspecifically, so the DNA sequence at the cut can be chosen to code for any two desired amino acids. This means that the DNA sequence can also be chosen equal to where a different restriction endonuclease cuts, for example *BanI* as in our case (**Fig. 2.1 a**). Because of this, it is easier to produce large seamless blocks where the repetitive sequence is continuous throughout the block. The second advantage of *BsaI* cutting next to its recognition site is that cloning can be sped up considerably, as can be seen in the following example. A vector containing for example (GAGAGAGE)₂ is linearized using *BsaI*. Separately, (GAGAGAGE)₂ is cut out of its vector using *BsaI* and *BanI*. This insert block is ligated into the linearized vector next to the other (GAGAGAGE)₂ block, creating (GAGAGAGE)₄. Because *BsaI* cuts next to its recognition site, the recognition site itself is not included in the junction of the two newly connected blocks but remains on one side of the whole new block. Therefore the block size is increased exponentially, namely by a factor of 2 per cloning step (**Fig. 2.2**), instead of linearly, namely by one basic block per cloning step.

Created by ligating the P2 DNA [11] into the *DraIII* digested adapter, the C block, also has *BsaI* and *BanI* restriction endonuclease sites (**Fig. 2.1 b**) that can be used to connect the C and S blocks in exactly the same way as the-silk blocks were enlarged (**Fig 2.2**). At the junctions of the blocks, the DNA sequence will code for glycine-alanine (**Fig 2.3**). After connection of the blocks, both restriction sites will be lost, at the junction, because glycine will be coded as in the *BanI* restriction site (**Fig. 2.1**) and alanine will be coded as in the *BsaI* restriction site (**Fig. 2.1**). In this way, we first combined C, S^E and S^H blocks and produced templates for diblocks (**Fig. 2.2**). Finally we combined the diblocks into the four genes encoding: CS^ES^EC, S^ECCS^E, CS^HS^HC and S^HCCS^H (**Fig. 2.2**), that were stably integrated in the *P.pastoris* genome. The reasons for using *P.pastoris* as an expression host are that repetitive genes are more stable in eukaryotes like *P.pastoris* than in prokaryotes like *E.coli*, and that *P.pastoris* has a proven record of good expression of repetitive genes [11, 13, 17].

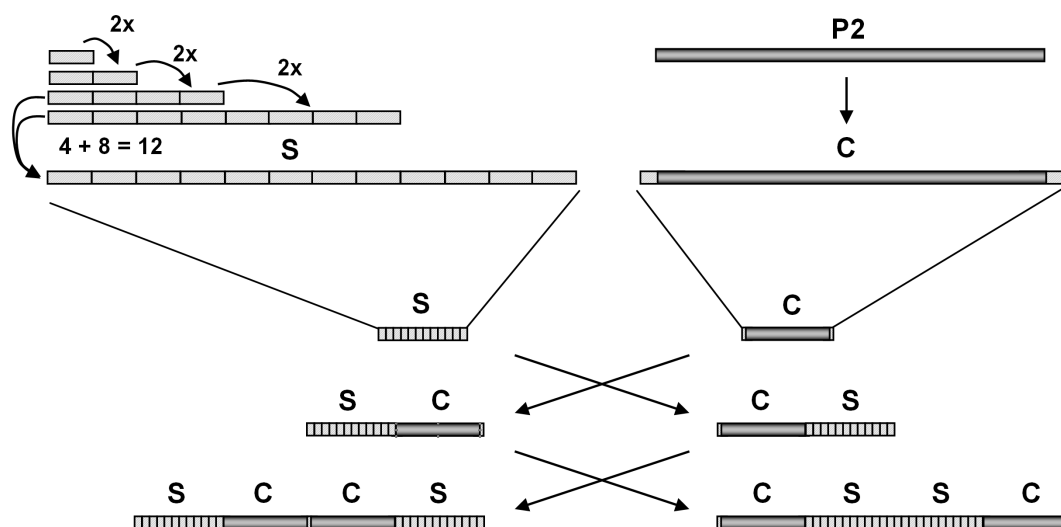


Figure 2.2 An overview of how the genes encoding the block copolymers were built by first creating different blocks of DNA and then combining them (in a cloning vector that is not depicted) using restriction and recursive ligation. On the top left, the silk-like repeat $(\text{GAGAGAGX})_2$, in which X is either glutamic acid or histidine, is elongated to 12 times its original length: $(\text{GAGAGAGX})_{24}$. On top the right, the P2 collagen-like sequence is compatibilized by adding restriction endonuclease sites to its ends that can be connected to the silk-like sequence. The silk-like block (S) and the collagen-like block (C) are then connected to form diblock encoding DNA sequences, which are further connected to effectively form triblock encoding genes.

2.3.2 Four protein block copolymers

Expression of the genes encoding the protein block copolymers resulted in the production of the amino acid sequences in **Figure 2.3**, from which the prepro secretion signal (highlighted first 89 amino acids) is cleaved off upon secretion, resulting in the four protein block copolymers that were purified from the cell-free broth using an easy, scalable procedure of selective precipitation. The conditions under which our products precipitated selectively from their mixture with other proteins and broth components depended only on the contents of S^{E} or S^{H} in the product, and not on block order. Starting from the amino acid sequence YVEFGLGA and ending on GA, both the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ molecules both have a molecular weight of 65750 Da, and both the $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ molecules have a molecular weight of 66135 Da. Together, the S blocks of one product contain 48 charged residues, either glutamic acid or histidine. Located in the C blocks, and the start sequence YVEFGLGA, there are an additional 19 glutamic acid and 12 lysine residues that can be charged. Except for the difference in histidine or glutamic acid, all four products have exactly the same amino acid composition.

$$CS^E S^E C$$

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPSNSTNNGLLFIINTIASIAAKEEGVSL
KREAEAYVEFGLGAGAPGEPGNPGSPGNQGGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGSQGNP
GKNGQPGSPGSQGSPPGNQGSPPGQPGNPGQPGGEQKPGNQGPAEPGNPGSPGNQGGQPGNKGSPGNPGQPGNEG
QPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGSPPGNQGSPPGQPGNPGQPGGEQKPGNQGPAEPGAG
GAGGEGAG
GAGGEGAG
AGEGAG
GEGAG
EGAG
GAG
SQGNPGKNGQPGSPGSQGSPPGNQGSPPGQPGNPGQPGGEQKPGNQGPAEPGNPGSPGNQGGQPGNKGSPGNPGQ
PGNEGQPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGSPPGNQGSPPGQPGNPGQPGGEQKPGNQGPA
GEGA

 $S^E C C S^E$ [illegible]
$$CS^H S^H C$$

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPSNSTNNGLLFIINTIASIAAKEEGVSL
KREA EAYVEFGLGAGAPGEPGNPGSPGNQGGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGSQGNP
GKNGQPGSPGSQGSQSPGNQGSQPGQPGNPGQPGGEQGKPGNQGPAGEPGNPGSPGNQGGQPGNKGSPGNPGQPGNEG
QPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPSGSPGNQGSQPGQPGNPGQPGGEQGKPGNQGPAGEGAG
AGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAG
GAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAGHGAGAGAG
AGHGAGAGAGAGHGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAG
GHGAGAGAGHGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAG
HGAGAGAGHGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGAGHGAGAGAGHG
GAGAGAGHGAGAGAGHGAGAGAPGEPGNPGSPGNQGGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQG
SQGNPGKNGQPGSPGSQGSQSPGNQGSQPGQPGNPGQPGGEQGKPGNQGPAGEPGNPGSPGNQGGQPGNKGSPGNPGQ
PGNEGQPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGSQSPGNQGSQPGQPGNPGQPGGEQGKPGNQGP
GEAG

$$S^H C C S^H$$
[illegible]

Figure 2.3 The four protein block copolymer products as encoded in the genes. The 89 amino acid long prepro secretion signal, that is cleaved off upon secretion of the protein, is highlighted gray, as are the glycine-alanine junctions between the blocks, encoded in the DNA template by the *Bsa*I and *Ban*I restriction sites. The four protein block copolymers that were purified from the fermentation medium, started with the amino acid sequence YVEFGLGA and ended on GA.

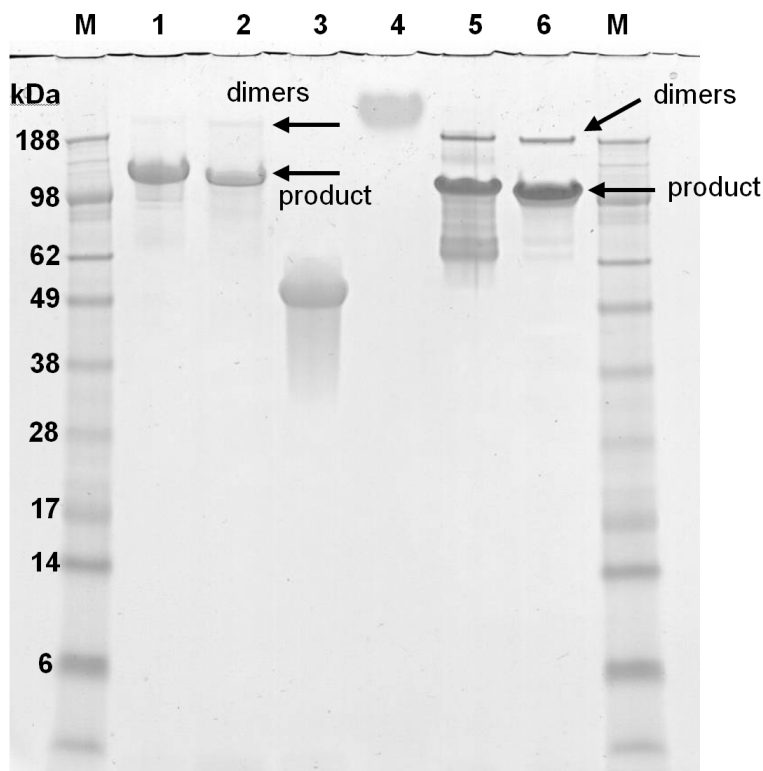


Figure 2.4 SDS-PAGE of purified products. Lanes M: molecular mass marker proteins; lane 1: $S^E CCS^E$; lane 2 $CS^E S^E C$; lane 3: pure $S^E S^E$ molecule; lane 4 pure CC molecule; lane 5: $CS^H S^H C$; lane 6: $S^H CCS^H$. Due to its extremely hydrophilic nature, CC (36.8 kDa) migrates to an extremely high apparent molecular mass [11], while $S^E S^E$ (28.2 kDa) migrate according to an apparent mass of about twice the true value. The migration of the triblocks is intermediate between that of CC and $S^E S^E$. Above the main band of the products there is a smaller band that might be caused by dimers.

2.3.3 Product identity and purity

After purification by selective precipitation, samples of, $CS^E S^E C$, $S^E CCS^E$, $CS^H S^H C$ and $S^H CCS^H$ were run on DSD page, blotted and N-terminally sequenced, showing that the products started with YVEFGL, confirming both the identity of our products and the intactness of their N-terminus. On SDS-PAGE, the products were compared to samples of pure $S^E S^E$ and CC (**Fig. 2.4**), which were kindly provided by M.W.T. Werten and F.A. de Wolf. Due to its extremely hydrophilic nature, CC (36.8 kDa) migrates to an extremely high apparent molecular mass [11], while $S^E S^E$ (28.2 kDa) migrates according to an apparent mass of about twice its true value. The migration of $CS^E S^E C$ and $S^E CCS^E$ is intermediate between that of CC and $S^E S^E$. The migrations of $CS^H S^H C$ and $S^H CCS^H$ are similar to those of $CS^E S^E C$ and $S^E CCS^E$. They have the C block in common, and differ only in the S block that contains either glutamic acid or histidine, so the migrational behavior of $S^E S^E$ is not due to its charge but is intrinsic to the $(GAGAGAGX)_n$ sequence with X being either E or H. In the SDS-

PAGE (**Fig. 2.4**) we observe a band, present for all products, of about twice the apparent molecular weight of the product. This extra band was not observed in the cell-free fermentation broth, but only in samples that were prepared from freeze-dried material. This extra band may be due to the formation of dimers of the product during the freeze drying process. The possible dimers were more pronounced for the S^H containing products than for the S^E containing products.

From conductivity measurements of dissolved protein, it appeared that our products still contained large amounts of salt, probably mostly ammonium sulphate that was co-precipitated with the protein in the acetone precipitation step. Extensive dialysis removed the salt and possibly low MW contaminants. After freeze-drying, it appeared that, depending on the product, between 20% and 40% of the original weight was retained. Conductivity measurements of dissolved protein now confirmed that salt was reduced to less than 4 wt% for S^E containing products (**Table 2.1**). For the S^H containing products salt content was even reduced to less than the detection limit of the method (0.5%) (**Table 2.1**).

After dialysis, amino acid content analysis (**Table 2.2**) revealed that 98 % or more of the protein content was indeed the intended product and that the remaining less than 2% was consistent with the amino acid composition of *P.pastoris* cell-free fermentation broth (**Table 2.1**). A contamination with merely 0.9-2% other proteins means that, in the SDS-PAGE (**Fig. 2.4**), the many minor bands observed under the main band of CS^HS^HC and S^HCCS^H must be our product, but in the form of different populations that migrate faster than the main band. An explanation for these additional bands, seen for the positively charged products could be that the positive charge of the self-assembling block is compensated by the negatively charged SDS, upon which, for a fraction of the S^H containing products, some protein folding occurs, leading to faster migration on SDS-PAGE.

	CS ^E S ^E C	S ^E CCS ^E	CS ^H S ^H C	S ^H CCS ^H
Product	90 %	90 %	93 %	95 %
<i>P.Pastoris</i> protein	1.1%	1.7%	0.9%	2.0%
(NH ₄) ₂ SO ₄	3 %	4 %	0 %	0 %
sugars	5.6%	4.3%	6.1%	3.2%

Table 2.1 Description of the purity of the products and the impurities found by chemical analysis.

Amino Acids		CS ^E S ^E C			S ^E CCS ^E		
		Expected %	Measured %	Measured %	Expected %	Measured %	Measured %
Ala	A	19.08	18.47	18.85	19.08	18.95	18.78
Asx	B = D + N	5.99	6.02	5.92	5.99	5.94	5.80
Phe	F	0.12	0.42	0.25	0.12	0.22	0.20
Gly	G	41.40	41.21	42.20	41.40	42.02	42.54
His	H	0.00	0.00	0.00	5.99	5.48	5.47
Ile	I	0.00	0.39	0.16	0.00	0.15	0.12
Lys	K	1.50	1.73	1.47	1.50	1.50	1.49
Leu	L	0.12	0.66	0.34	0.12	0.35	0.30
Pro	P	11.22	10.79	11.00	11.22	11.06	11.17
Arg	R	0.00	0.00	0.00	0.00	0.00	0.00
Ser	S	3.99	3.20	3.22	3.99	3.17	3.20
Thr	T	0.00	0.58	0.31	0.00	0.35	0.24
Val	V	0.12	0.74	0.39	0.12	0.38	0.29
Tyr	Y	0.12	0.39	0.26	0.12	0.24	0.22
Glx	Z = E + Q	16.33	15.40	15.63	10.35	10.19	10.18
		100	100	100	100	100	100

Table 2.2 Result of the amino acid content analysis in percentages of the total molar amino acid content. In the first box, the amino acids. In the second box, the expected and measured content of amino acids for the CS^ES^EC and S^ECCS^E respectively. In the third box, the expected and measured content of amino acids for the CS^HS^HC and S^HCCS^H respectively.

Sugar assay revealed that the products still contained between 3.2 % and 5.6 % of sugars (Table 2.1), probably polysaccharides, because small sugars should have been removed during dialysis. Each fermentation took about 4 days and yielded approximately 1.5 l of cell-free broth. Based on the weight and purity of the recovered products, the total amounts of product recovered from the broth were 1.7 g, 1.0 g, 0.8 g, and 1.5 g for CS^ES^EC, S^ECCS^E, CS^HS^HC, and S^HCCS^H respectively, making the recovered product yields from the fermentation broth: 1.13 g/l, 0.66 g/l, 0.53 g/l, and 1 g/l respectively. These are amongst the highest yields for secreted heterologous protein published to date [13, 21, 22].

2.3.4 Product aggregation forming clear gels

Based on earlier work [23], we expected that upon rendering the silk blocks uncharged by shifting the pH, they would aggregate. To our satisfaction, CS^ES^EC and S^ECCS^E indeed formed transparent gels (Fig. 2.5) at pH2 and at an exceptionally low concentration of 0.9 g/l. In contrast, 0.9 g/l CS^HS^HC and S^HCCS^H did at first not seem to aggregate at pH 12, but at 4.5 g/l they formed transparent gels that could support their own weight (Fig. 2.5). Probably aggregates were formed at 0.9 g/l but these did not form a robust network like CS^ES^EC and

$S^E CCS^E$ did. So, they obviously formed self-assembled networks with dimensions smaller than the wavelength of visible light, otherwise the gels would not be transparent. As intended, the combination in one molecule of aggregating S blocks and hydrophilic C blocks that stop aggregation led to nano-sized objects forming a network.

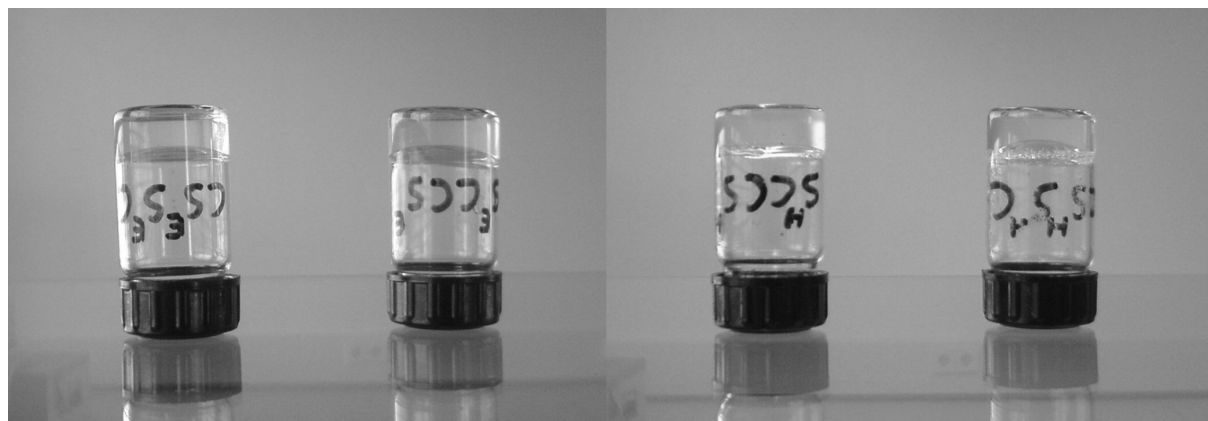


Figure 2.5 From left to right respectively: vials containing 0.9 g/l gels of $CS^E S^E C$ and $S^E CCS^E$, formed at pH2, and of 4.5 g/l $S^H CCS^H$ and $CS^H S^H C$, formed at pH12. The gels were formed on the bottom of the vials, after which they were turned upside down. the photos were taken after half an hour.

2.4 Conclusion

Four genes of seamless blocks of highly repetitive DNA were produced using the cloning approach described above. The genes were well expressed in *P.pastoris* and recovered yields of the encoded protein block copolymers were around 1 g/l of cell-free fermentation broth. The total amounts recovered were 1.7 g, 1.0 g, 0.8 g, and 1.5 g for $CS^E S^E C$, $S^E CCS^E$, $CS^H S^H C$, and $S^H CCS^H$ respectively, which will be sufficient for a variety of experiments. After purification, the products were intact, and at least 90% pure, with the contaminants being mainly some salt and some polysaccharides. Having S^E or S^H in the product did not have much influence on the migrational behavior of the whole molecule on SDS-PAGE. Therefore the anomalously slow migration of $S^E S^E$ may not be due to the glutamic acid residues and their negative charge but may be intrinsic to the silk-like block repeat. On the SDS-PAGE we also observed bands above the main product bands that might be dimers of the products caused by the freeze-drying process. They were more pronounced for the S^H containing products than for the S^E containing products. Some fractions of the S^H containing products, possibly folded under influence of the oppositely charged SDS migrated faster than the main band. When the pH was adjusted appropriately to reduce the charge of the silk-like blocks, all products formed transparent gels. The S^E containing products did so at an

exceptionally low concentration of 0.9 g/l, but S^H containing products needed higher concentrations to produce a gel. In the following chapters we will investigate the structures responsible for this gelling behavior.

Acknowledgements

The authors thank Antoine Moers for help with, and Cheng Tao for his contributions to fermentations and protein purification.

References

- 1 Chen, C. S., Mrksich, M., Huang, S., Whitesides, G. M. and Ingber, D. E. (1997) Geometric control of cell life and death. *Science* **276**, 1425-1428
- 2 Chitkara, D., Shikanov, A., Kumar, N. and Domb, A. J. (2006) Biodegradable injectable in situ depot-forming drug delivery systems. *Macromolecular Bioscience* **6**, 977-990
- 3 Harrington, D. A., Cheng, E. Y., Guler, M. O., Lee, L. K., Donovan, J. L., Claussen, R. C. and Stupp, S. I. (2006) Branched peptide-amphiphiles as self-assembling coatings for tissue engineering scaffolds. *Journal Of Biomedical Materials Research Part A* **78A**, 157-167
- 4 Krausch, G. and Magerle, R. (2002) Nanostructured thin films via self-assembly of block copolymers. *Advanced Materials* **14**, 1579-1583
- 5 Muthukumar, M., Ober, C. K. and Thomas, E. L. (1997) Competing interactions and levels of ordering in self-organizing polymeric materials. *Science* **277**, 1225-1232
- 6 Park, M., Harrison, C., Chaikin, P. M., Register, R. A. and Adamson, D. H. (1997) Block copolymer lithography: Periodic arrays of similar to 10(11) holes in 1 square centimeter. *Science* **276**, 1401-1404
- 7 Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A. and Stupp, S. I. (2004) Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science* **303**, 1352-1355
- 8 Woolfson, D. N. and Ryadnov, M. G. (2006) Peptide-based fibrous biomaterials: some things old, new and borrowed. *Current Opinion In Chemical Biology* **10**, 559-567
- 9 Krejchi, M. T., Cooper, S. J., Deguchi, Y., Atkins, E. D. T., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Crystal structures of chain-folded antiparallel beta-sheet assemblies from sequence-designed periodic polypeptides. *Macromolecules* **30**, 5012-5024
- 10 Parkhe, A. D., Cooper, S. J., Atkins, E. D. T., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1998) Effect of local sequence inversions on the crystalline antiparallel beta-sheet lamellar structures of periodic polypeptides: implications for chain-folding. *International Journal of Biological Macromolecules* **23**, 251-258
- 11 Werten, M. W. T., Wisselink, W. H., van den Bosch, T. J. J., de Bruin, E. C. and de Wolf, F. A. (2001) Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Engineering* **14**, 447-454
- 12 Rozkiewicz, D. I., Kraan, Y., Werten, M. W. T., de Wolf, F. A., Subramaniam, V., Ravoo, B. J. and Reinhoudt, D. N. (2006) Covalent microcontact printing of proteins for cell patterning. *Chemistry-A European Journal* **12**, 6290-6297
- 13 Werten, M. W. T., Van den Bosch, T. J., Wind, R. D., Mooibroek, H. and De Wolf, F. A. (1999) High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**, 1087-1096
- 14 Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) *Pichia-Pastoris* as a Host System for Transformations. *Molecular and Cellular Biology* **5**, 3376-3385
- 15 Fahnstock, S. R. and Irwin, S. L. (1997) Synthetic spider dragline silk proteins and their production in *Escherichia coli*. *Applied Microbiology And Biotechnology* **47**, 23-32

- 16 Cregg, J. M., Cereghino, J. L., Shi, J. Y. and Higgins, D. R. (2000) Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology* **16**, 23-52
- 17 Fahnestock, S. R. and Bedzyk, L. A. (1997) Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Applied Microbiology And Biotechnology* **47**, 33-39
- 18 Chambers, S. P., Prior, S. E., Barstow, D. A. and Minton, N. P. (1988) The Pmtl Nic-Cloning Vectors .1. Improved Puc Polylinker Regions To Facilitate The Use Of Sonicated Dna For Nucleotide Sequencing. *Gene* **68**, 139-149
- 19
- 20 Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. and Smith, F. (1956) Colorimetric Method For Determination Of Sugars And Related Substances. *Analytical Chemistry* **28**, 350-356
- 21 Cereghino, G. P. L., Cereghino, J. L., Ilgen, C. and Cregg, J. M. (2002) Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Current Opinion In Biotechnology* **13**, 329-332
- 22 Jana, S. and Deb, J. K. (2005) Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Applied Microbiology and Biotechnology* **67**, 289-298
- 23 Werten, M. W. T., Moers, A. P. H. A., Vong, T., Zuilhof, H., van Hest, J. C. M. and de Wolf, F. A. (2008) Biosynthesis of an Amphiphilic Silk-Like Polymer. *Biomacromolecules*

3

Self-assembled and co-assembled structures of silk-collagen-like block copolymers

(with themselves, with oppositely charged polyelectrolytes, and with each other)

A.A. Martens, M. Schor, P.G. Bolhuis, G. Portale, M. Neeleman, R.J. de Vries, Y. Yan,
F. Li, G. Eggink, F.A. de Wolf, M.A. Cohen Stuart

Abstract

The self-assembled structure of a protein block copolymer, depends on its amino acid sequence, but also on several other factors in the molecular environment like solvent, pH, or other molecules with which it might interact or even co-assemble. We describe the self-assembled, and co-assembled structures of a set of four complementary, oppositely charged protein block copolymers with themselves, with oppositely charged polyelectrolytes and with each other. The four protein triblock copolymers under investigation are the 66 kDa $CS^E S^E C$, $S^E CCS^E$, $CS^H S^H C$ and $S^H CCS^H$. At neutral pH, the histidine containing S^H block and the glutamic acid containing S^E block are positively and negatively charged, respectively, the theoretical pKa being 7 and 4.3 respectively. Several self- and co-assembled structures were examined: 1-4) the four pH- induced self-assemblies, 5-7) $CS^H S^H C$ with poly acrylic acid (PAA), or with the metal bis-ligand supramolecular polymer ($Zn-L_2(EO)_4$), or with the (conducting) polythiophene (POWT) that was chemically modified to be zwitterionic, 8) $CS^E S^E C$ with POWT, 9-12) the four possible mixtures of the four different protein block copolymers. Except for the mixtures of protein polymers, which seemed to form kinetically trapped molecular networks, the self- and co-assemblies formed well defined μm long nanoribbons with a hydrophilic C block corona. The core structure, depended on the protein block copolymer used, and on the mode of charge compensation like pH and/or the type of polyelectrolyte used. Interesting features are: the unusual β -roll, predicted with MD modeling for the S^E block in the self-assembled ribbon core at low pH, the CC middle blocks of $S^E CCS^E$ and $S^H CCS^H$ forming loops, analogous to flower-like micelles, and the templating of the conductive polymer POWT onto nanoribbons that might have applications as nanowire.

3.1 Introduction

Developments in molecular biology have facilitated production of designed proteins that can be regarded as block copolymers. These biosynthetic protein polymers present opportunities to access nano-structures previously unobtainable with classical (block-co)polymers. Therefore the field of protein polymer science is expanding rapidly. The design of protein polymers is often nature-inspired because nature has already provided a library of natural amino acid sequences with a structure-forming function [1]. The reasons for protein polymers to self-assemble into such highly defined nanostructures are that they have an exactly defined domain structure in which each domain has an exactly defined length and repetitive monomer sequence. These are features typically lacking in classical polymers. Not only structure, but also function, can be given to the formed nanostructures, by incorporation of additional blocks with functional amino acid sequences, ultimately leading to high value applications like regenerative medicine, pharmaceuticals, self-healing coatings, chemomechanical fibers, sensors, smart packaging, and others [2-8].

Stimulus-responsive formation of nanostructures is a feature typically obtained with block copolymers. In chemically synthesized triblock copolymers, the order of the blocks normally determines the supramolecular organization. A polymer with soluble end blocks and an insoluble middle block usually assembles into micelles, vesicles or lamellar structures [9]. Conversely, a polymer with insoluble end blocks and a soluble middle block can form molecular networks and gels [10]. The question is whether or not the typical behavior of chemical triblock copolymers is also found in self-assembling protein polymers with a triblock structure.

In the present work we describe the self-assembled and co-assembled structures of the four triblock copolymers described in **Chapter 2** $CS^E S^E C$, $S^E CCS^E$, $CS^H S^H C$, and $S^H CCS^H$. We will further investigate the morphology of the self-assembled fibrils and the different conformational structures that the protein polymers may take. These depend on several factors like: the type of S block used in the molecule (S^E or S^H), the arrangement (CSSC or SCCS), and the pH. When co-assembling with an oppositely charged polyelectrolyte at moderate pH, different types of polyelectrolyte seem to lead to different final structures of the protein fibrils. Co-assembly of the four products with oppositely charged polyelectrolytes, that would normally not form linear aggregates gives us opportunities to nevertheless produce fibrils with

such polymers. One example could be co-assembly with conductive polymers that might lead to applications as nanowires. All in all, we present many different ways to manipulate the nanostructures formed by our protein polymers, which will be useful when tailoring the materials formed by them to an application. The results and discussion are presented in three sections: section I, in which the structures formed by S^ECCS^E and CS^ES^EC at low pH are described, section II, in which the structures formed by S^HCCS^H and CS^HS^HC at high pH are described, and section III, in which the structures are described formed by the protein block copolymers in combination with various oppositely charged polyelectrolytes and in combination with each other.

3.2 Materials and methods

3.2.1 Protein block copolymers

The synthesis of the protein block copolymers CS^ES^EC , S^ECCS^E , CS^HS^HC , S^HCCS^H and their primary amino acid sequences have been described previously in **Chapter 2**. Unless mentioned otherwise, stock solutions contained 1 g/l of product in 1 mM NaOH or HCl. CS^ES^EC and S^ECCS^E were prepared in NaOH and CS^HS^HC and S^HCCS^H were prepared in HCl. Both Pure CC block [11] and pure S^ES^E polymer were taken from a batch that was produced before [12].

3.2.2 Polyelectrolytes

Three different polyelectrolytes were used to compensate the charge of the protein polymers. Zwitterionic poly(3-[(S)-5-amino-5-carboxyl-3-oxapentyl]-2,5-thiophenylene hydrochloride) (POWT) (**Fig. 3.1 a**) of 15 monomers long, with an equivalent of 15 either positive or negative charges per molecule was synthesized as described before [13]. Polyacrylic acid (PAA) (**Fig. 3.1 b**) with a molecular weight of 3000, equivalent to 42 monomers and negative charges per molecule, was purchased from Polymer Source (Canada) and neutralized with NaOH, to form sodium polyacrylate. Negatively charged zinc(II)-1,11-bis(2,6-dicarboxypyridin-4-yloxy)-3,6,9-trioxaundecane metallo-supramolecules $(Zn-L_2(EO)_4)^{2-}$ complex was synthesized as described before [14] (**Fig. 3.2**). Products were dissolved in water or buffer and pH adjusted with HCl or NaOH.

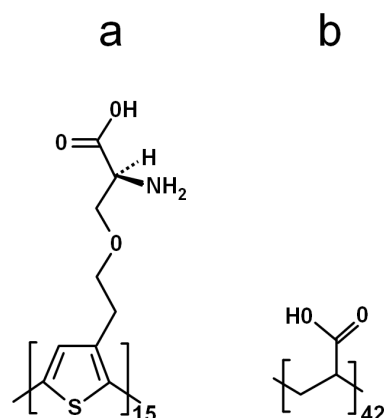


Figure 3.1 Molecular structures of polyelectrolytes used to co-assemble with protein block copolymers. (a) POWT: poly(3-[(S)-5-amino-5-carboxyl-3-oxapentyl]-2,5-thiophenylene hydrochloride) and (b) PAA: polyacrylic acid.

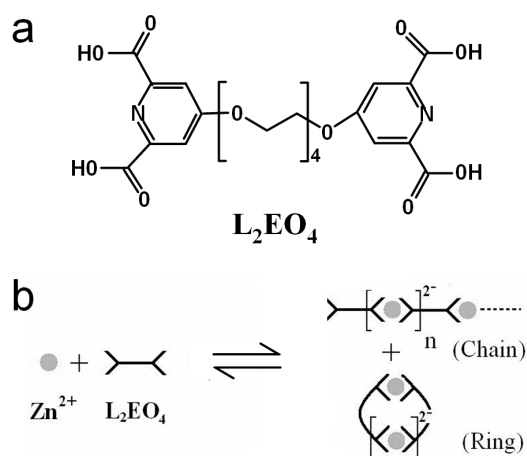


Figure 3.2 Molecular structure of (a) Bisligand $L_2(EO)_4^{4-}$: 1,11-bis(2,6-dicarboxypyridin-4yloxy)-3,6,9-trioxaundecane, which can coordinate Zn^{2+} to form (b) $(Zn-L_2(EO)_4)_n$ complex with an overall charge of -2 per $Zn-L_2(EO)_4^{2-}$ unit. The complex can form rings of 2 units at low concentrations and linear coordination polymers at high concentrations.

3.2.3 Microscopy of pH induced $CS^E S^E C$ and $S^E CCS^E$ filaments

A stock solution of $CS^E S^E C$ or $S^E CCS^E$ was diluted with 10 mM HCl to a final product concentration of 0.1 g/l and left overnight at room temperature to allow supramolecular assembly. The resulting dispersion was used to prepare samples for AFM as well as TEM and cryo-TEM.

For AFM imaging of supramolecular structures (**Fig. 3.3 a,b**), a piece of clean hydrophilic silica wafer was dipped into the $CS^E S^E C$ or $S^E CCS^E$ dispersion, rinsed shortly with dematerialized water, dried under a stream of nitrogen and analyzed using a Digital Instruments NanoScope III in tapping mode.

For TEM imaging (**Fig. 3.3 c,d**), a grid covered with Formvar film was first contacted with a drop of the $CS^E S^E C$ or $S^E CCS^E$ dispersion, then with a drop of 20 g/l uranyl acetate (as a staining agent) in water (pH 3.7), subsequently air-dried and viewed using a JEOL JEM 1200 EX II microscope operated at 80 kV. Digital images were recorded with KeenView/iTEM (SIS, Munster, Germany).

For cryo-TEM imaging of pH induced filaments (**Fig. 3.4**), an aliquot of 3 μ l sample solution was pipetted onto a glow-discharged Quantifoil R2/2 copper grid 200 mesh in the environmental chamber of a Vitrobot at a relative humidity of 100 %. The sample was blotted

once during 0.5 s and rapidly plunged into liquid ethane. The grid was subsequently transferred to a Gatan cryoholder Model 626 and viewed in a Philips Tecnai12 TEM equipped with a Biotwin-lens and a LaB6 filament. Images were recorded with a SIS Megaview II CCD-camera and processed with AnalySIS software.

3.2.4 Microscopy of pH induced $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ filaments

A stock solution of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ or $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ was diluted with 10 mM NaOH to a final product concentration of 0.1 g/l and left overnight at room temperature to allow supramolecular assembly. The resulting dispersion was used to prepare samples for AFM as well as TEM and cryo-TEM.

For AFM imaging of supramolecular structures (**Fig. 3.13 a,b**), a piece of clean hydrophilic silica wafer was dipped into the $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ or $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ dispersion, rinsed shortly with 10⁻⁵ M NaOH (pH9), dried under a stream of nitrogen and analyzed using a Digital Instruments NanoScope III in tapping mode.

TEM (**Fig. 3.13 c,d**), and cryo-TEM (**Fig. 3.14**) imaging, was done as described above for $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$.

3.2.5 Microscopy of POWT induced filaments of both $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$

A dispersion of co-assembled $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and POWT was prepared by mixing a 10 g/l solution of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ in 10 mM HCl with a 1 g/l POWT solution, and diluting to final concentrations of 0.3 g/l $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and 0.03 g/l POWT, the concentration ratio of which is equivalent to a charge ratio of protein / POWT = 1.5. A similar dispersion was prepared with a 10 g/l $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ solution in 10 mM NaOH at the same protein/POWT charge ratio.

For AFM imaging of the co-assemblies (**Fig. 3.17 a,d**), a few drops of co-assembly dispersion was deposited on freshly cleaved mica, and dried overnight at room temperature. Imaging in tapping-mode was carried out using a JSPM-5400 scanning probe microscope from JEOL (Europe, BV) and NSC35/AIBS ultra sharp cantilevers (MikroMasch, Europe).

For TEM imaging (**Fig. 3.17 b,c,e,f**), a grid covered with Formvar film was first contacted with a drop of co-assembly dispersion, then with a drop of 20 g/l uranyl acetate (as a staining agent) in water (pH 3.7), subsequently air-dried and viewed using a JEOL JEM 1200 EX II microscope operated at 80 kV. Digital images were recorded with KeenView/iTEM (SIS, Munster, Germany).

3.2.6 Microscopy of CS^HS^HC filaments induced by PAA or Zn-L2EO4

A dispersion of co-assembled CS^HS^HC and Zn-L₂(EO)₄ was prepared as follows [15]. A solution of Zn-L₂(EO)₄ complex was added to a CS^HS^HC stock solution until a charge ratio of 1:1 was reached. The sample was then diluted to a to the final concentration of 0.65 g/l CS^HS^HC. A dispersion of co-assembled CS^HS^HC and PAA was prepared similarly.

For cryo-TEM (**Fig. 3.18 a-d**), A few microliters of samples were placed on a bare copper TEM grid (Plano, 600 mesh), and the excess liquid was removed with filter paper. This sample was cryo-fixed by rapidly immersing into liquid ethane cooled to -170 to -180 °C in a cryo-box (Carl Zeiss NTS GmbH). The specimen was inserted into a cryo-transfer holder (CT3500, Gatan, Munich, Germany) and transferred to a Zeiss EM922 EFTEM (Zeiss NTS GmbH, Oberkochen, Germany). Examinations were carried out at temperatures around -180 °C. The TEM was operated at an acceleration voltage of 200 kV. Zero-loss filtered images were taken under reduced dose conditions (500-2000 e/nm²). All images were recorded digitally by a bottom-mounted CCD camera system (UltraScan 1000, Gatan) and processed with a digital imaging processing system (Digital Micrograph 3.9 for GMS 1.4, Gatan).

For cryo-TEM (**Fig. 3.18 a-d**), filaments were imaged as described above for CS^ES^EC and S^ECCS^E.

3.2.7 Microscopy S^E containing products mixed with S^H containing products

Solutions of 0.1 g/l of CS^ES^EC or S^ECCS^E in 0.1 mM NaOH were mixed in a ratio of 1:1 with solutions of 0.1 g/l of CS^HS^HC or S^HCCS^H in 0.1 mM HCl and imaged using TEM (**Fig. 3.22 a-d**), as described above for CS^ES^EC and S^ECCS^E.

3.2.8 CD spectroscopy of the different products under different conditions

All CD spectra were recorded between 190 and 260 nm on a Jasco J-715 spectropolarimeter at 21°C, with a resolution of 0.2 nm and a scanning speed of 1 nm/s. The spectra shown are the average of 15 recorded spectra. All samples were transferred to a 1 mm quartz cuvette directly after preparation, after which the pH of the remaining fluid was verified. Samples were allowed to age in their cuvette during which CD measurements were taken.

Samples of CS^ES^EC or S^ECCS^E (**Fig. 3.5, 3.6, 3.7**) were prepared at different pH by mixing a stock solution in with water and 1 M HCl, resulting in final polymer concentrations of 0.1 g/l at pH 7, and 0.09 g/l at pH 2. A sample of 0.1 g/l pure S^ES^E at pH 7 was prepared

similarly. Samples of pure CC were prepared at different pH by diluting a 1 g/l CC stock solution in demineralized water and adding 1 M HCl or 1 M NaOH, resulting in final polymer concentrations of 0.1 g/l at pH 7, and 0.09 g/l at various pH.

Samples of $CS^H S^H C$ or $S^H C C S^H$ (**Fig. 3.15**) were prepared at pH 7, and at pH 12 by mixing a $CS^H S^H C$ or $S^H C C S^H$ stock solution with water and 1 M NaOH, resulting in a final polymer concentration of 0.1 g/l.

Samples of 0.1 g/l $CS^H S^H C$ under different conditions (**Fig. 3.20**) were prepared from stock solutions, all containing 20 mM PIPES buffer pH 5.4. Also the 1 g/l $CS^H S^H C$ stock solution contained 20 mM PIPES buffer. A sample of 0.1 g/l $CS^H S^H C$ was prepared by diluting the stock with 20 mM PIPES buffer. The sample of 0.1 g/l $CS^H S^H C$ at pH 11 was prepared by adding 1M NaOH and water to the stock solution. A sample of $CS^H S^H C$ with $Zn-L_2(EO)_4$ (pH5.4) at a charge ratio of 1:1 was prepared by mixing stock solutions of both components, and diluting with 20 mM PIPES buffer to 0.1 g/l of protein. A sample of $CS^H S^H C$ with PAA (pH5.4) at a charge ratio of 1:1 was prepared similarly.

To prepare samples of mixed protein block copolymer products (**Fig. 3.21**), S^E containing and S^H containing polymers were mixed in a ratio of 1:1. Components (if used) were always added to the mix in the following order: water or 5 mM phosphate buffer, 1M NaCl, stock solution of $CS^E S^E C$ or $S^E C C S^E$, stock solution of $CS^H S^H C$ or $S^H C C S^H$. The unbuffered samples contained a total of 0.1 g/l protein. Final concentrations in the buffered samples were 0.1 g/l total protein and 4.5 mM phosphate buffer. Final concentrations in the salty samples were 0.1 g/l total protein, 4.25 mM phosphate buffer and 50 mM NaCl.

3.2.9 MD modeling of the S^E block

All MD simulations were performed with the GROMACS molecular simulation package [16]. The OPLSAA force field [17] was used for the protein in combination with the SPC water model [16]. Bonds were constrained by LINCS [18], allowing for a 2 fs time step. Long-range electrostatics were treated with fast Particle-Mesh Ewald [19] with a grid spacing of 1.1 Å. A cut-off range of 1.1 Å was used for the van der Waals interactions. After an energy minimization run and a short peptide-restraint run, an equilibration run of 1ns was performed at constant pressure of 1 bar and a temperature of 298 K using Parrinello-Rahman coupling and a Nosé-Hoover thermostat [20]. Subsequent MD simulations were performed with the same settings.

Starting structures were created with the MOLMOL software [21]. They had the sequence GAGAGAGEGAGAGA and were created as hairpins with ψ and ϕ angles based on 1SLK or 2SLK structures [22] and a turn [23] containing the middle glutamic acid. Similarly, the resulting stable structure: a twisted hairpin (**Fig. 3.8 a**), was used to create the starting structure of the sequence E(GAGAGAGE)₁₀. Similarly, the resulting β -roll structure (**Fig. 3.8 b**) was duplicated and stacked manually to simulate a possible stack (**Fig. 3.8 c**). They were stacked manually and energy was minimized. The PDB output files of the equilibrated structures were imported into the YASARA molecular graphics software packet [24] which was only used to edit the appearance for illustration purposes (**Fig. 3.8, 3.9**).

3.2.10 Size determination of the CC block with dynamic light scattering (DLS)

A 10 g/l CC stock solution was prepared by dissolving pure freeze-dried CC molecule in water. A sample at pH 1.5, of 0.57 g/l CC (which is the same concentration of C block as in a 1 g/l CS^ES^EC or S^ECCS^E solution) was prepared by diluting the stock and acidifying with HCl. Measurements were taken at a fixed angle of 90° in a light scattering setup equipped with a DPSS laser emitting vertically polarized light at a wavelength of 532.0 nm, an ALV-5000 multiple τ digital correlator and a Photo Multiplier Tube (PMT) detector. We did 5 runs of 5 minutes at 20°C. We estimated the radius twice, using two different methods. The first radius determination was based on the first cumulant of the second order cumulant fit to the 5 runs. The 5 obtained radii were averaged. The second radius determination was done by averaging the 5 autocorrelation functions from the 5 runs to one curve and averaging the radius from 5 contin fits to this one curve.

3.2.11 Small angle x-ray scattering (SAXS) measurements of supramolecular fibrils

A 20 g/l gel of CS^ES^EC at pH 1.5 was prepared, by first dissolving 0.1 g of freeze-dried material in 1 ml 100 mM NaOH and 3.742 ml of water, and then adding 0.258 ml 1M HCl to lower the pH and cause gelling. A 20 g/l gel of S^ECCS^E at pH 1.5 was prepared similarly. A 20 g/l gel of CS^HS^HC at pH 12 was prepared by first dissolving 0.1 g of freeze-dried material in 1 ml 100 mM HCl and 3.742 ml of water and then adding 0.258 ml 1M NaOH to increase the pH and cause gelling. A 20 g/l gel of S^HCCS^H at pH 12 was prepared similarly. The samples were allowed to gel for several days in glass vials. Core samples were taken by pushing 2mm kapton capillaries into the gels.

SAXS measurements (**Fig. 3.11, 3.16**) were performed on the samples contained in the 2mm kapton capillaries at the ESRF in Grenoble (France) on the BM26B Dutch-Belgian beamline (DUBBLE) with an X-ray photon energy of 13 keV and a sample to detector distance of 4040mm [25]. The acquisition time per image was 200 seconds, and the explored d range was 52.2 to 2.5 nm. Information about the scattering objects shape and dimensions were obtained via Guinier analysis.

3.3 Results and discussion Section I

pH induced self-assembly of the negatively charged $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$

3.3.1 Supramolecular ribbons of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ at low pH

This work describes for the first time high molecular weight silk-collagen triblock copolymers with two complementary arrangements of self-assembling and non-assembling inner and outer blocks, both forming gels.

To investigate the nano-structures responsible for the gelling of the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ solutions, atomic force microscopy (AFM), transmission electron microscopy (TEM) and cryogenic (cryo)-TEM micrographs of material adsorbed from dilute (0.1 g/l) $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ or $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ dispersions at pH 2 were produced (**Fig. 3.3, 3.4**).

In the AFM images, we observed similar, very long supramolecular fibrils of both $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ (**Fig. 3.3 a,b**) that were adsorbed to the silica support from dilute dispersions. For both polymers, the same fibril height of 1.5-2.0 nm was detected, and double that height at intersections of two fibrils. For both polymers, the same tip-convoluted, hence apparent width of several nanometers was deduced. However, dimensions of adsorbed and dried fibrils may, because of conformational changes, not be exactly the same as dimensions of fibrils in aqueous solution. These dimensions closely resemble those reported for PEO-conjugated poly(GAGAGAGE) [26], crystallized from organic solvents, and (non-conjugated) poly-(GAGAGAGX) with Tyr, Glu, His and Lys residues in consecutive X-positions (the so-called YEHK molecule), deposited from water [27]. The presence of supramolecular fibrils was confirmed by transmission electron microscopy (TEM) of negatively stained preparations (**Fig. 3.3 c,d**), which showed that both $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ fibrils have an apparent width of 5 to 6 nm, similar to YEHK fibrils [27]. From AFM and TEM images we estimate that at least 90 % of the fibrils are longer than 10 μm .

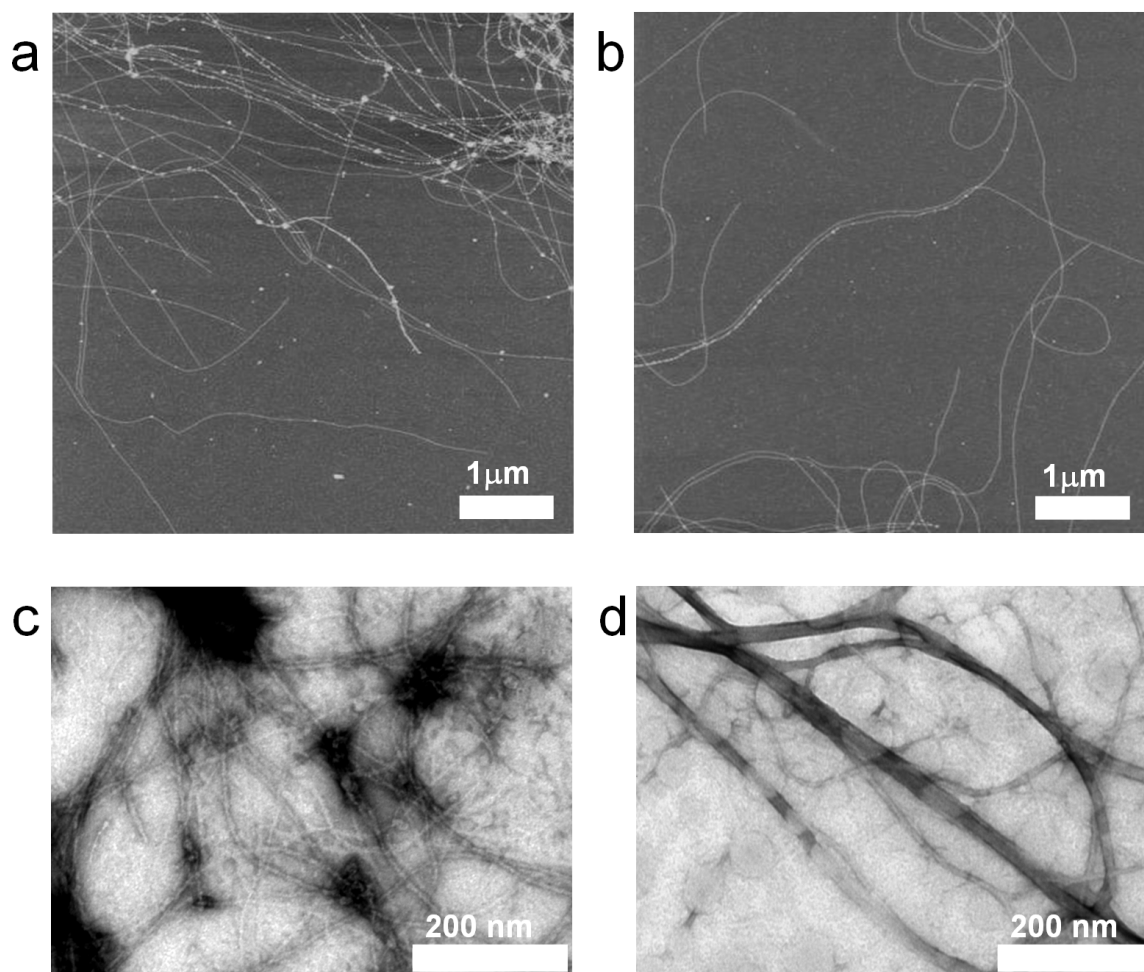


Figure 3.3 Micrographs of supramolecular $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ fibrils formed at pH 2. Fibrils in all images were deposited from a 0.1 g/l protein solution. (a, b) Atomic force micrographs of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ fibrils, respectively, deposited on silica. Higher objects are lighter in color. (c, d) TEM images of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ fibrils, respectively, deposited on Formvar.

It took several hours before the fibrils were formed, because they were not yet seen in AFM and TEM samples prepared shortly after acidification. This observation, and also the occurrence of fibrils in (cryo)-TEM images of solutions that were frozen after overnight incubation at pH 2 (**Fig. 3.4**), reveal that the formation of fibrils was not merely due to the process of drying the acidified solutions for subsequent microscopic analysis. In AFM images, the doubled apparent fibril thickness at intersections of the fibrils is in support of the fibrils forming in solution before being deposited on the silica.

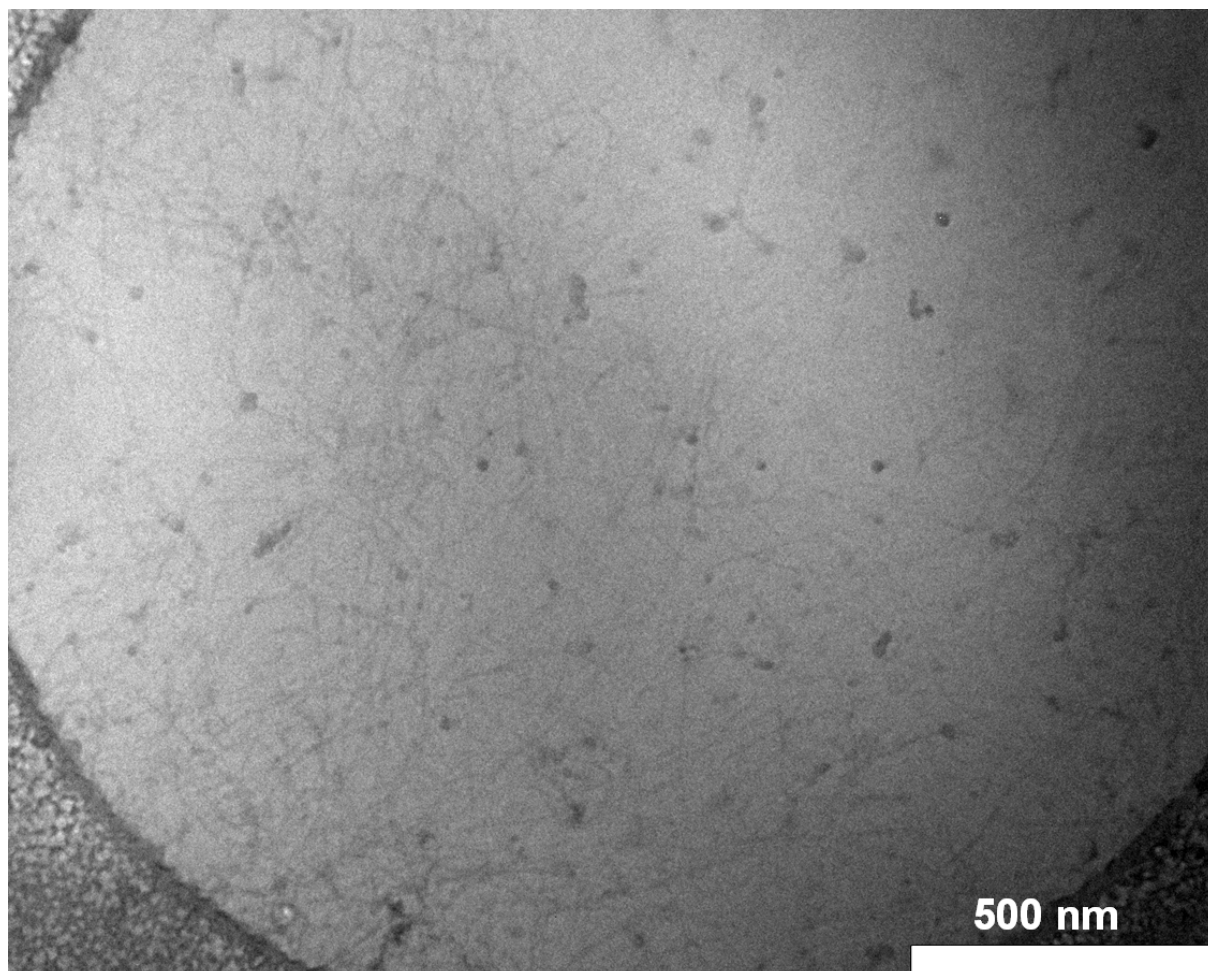


Figure 3.4 Cryo-TEM micrograph of $CS^E S^E C$ fibrils in copper grid. Fibrils were deposited from a 0.1 g/l protein solution.

The essential discriminative feature of $CS^E S^E C$ and $S^E CCS^E$ with respect to chemically synthesized block copolymers is their exactly defined monomer sequence and block length, in combination with the chiral nature of all monomer constituents. If the middle blocks in typical chemical polymers are insoluble, they form an amorphous phase inside nano-sized spherical or elongated particles. These are surrounded by a corona of soluble end blocks, stabilizing the particles against aggregation [9]. Conversely, if the end blocks are insoluble, very different structures such as networks, gels or flower-like micelles are usually obtained, because the two insoluble end blocks are separated by the mid block and will thus often reside in separate domains [10]. In contrast, the entirely biosynthetic $CS^E S^E C$ and $S^E CCS^E$ do not partition into an amorphous phase and both polymers form well-defined crystalline fibers with a morphology that is independent of the block order. This is a unique property for which we find no counterpart in the literature.

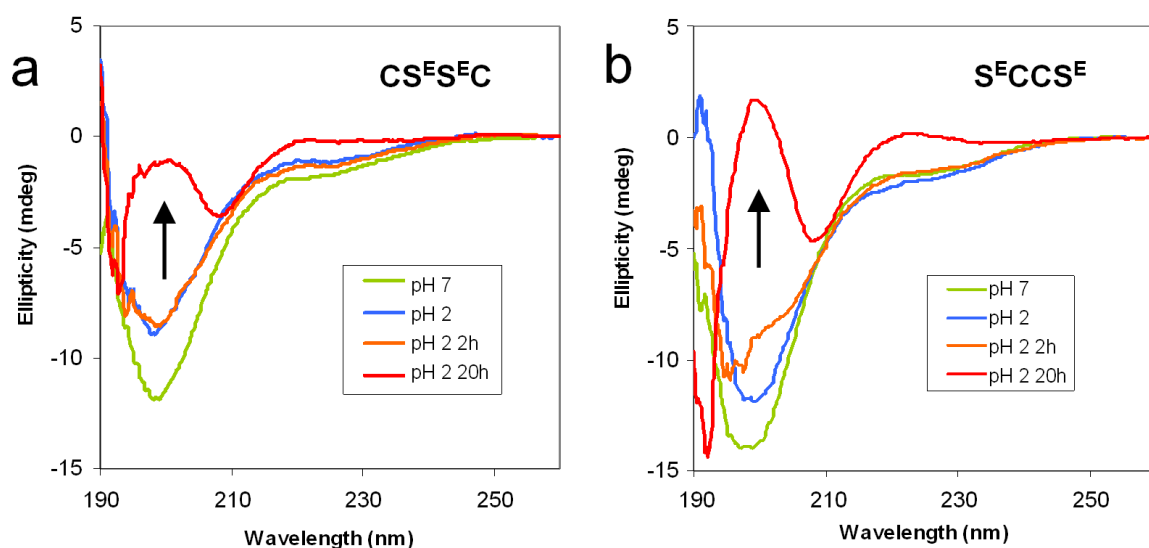


Figure 3.5 CD spectral changes reflecting structural changes induced by shifting the pH from 7 to 2 of: (a) 0.1 g/l $CS^E S^E C$ and (b) 0.1 g/l $S^E CCS^E$.

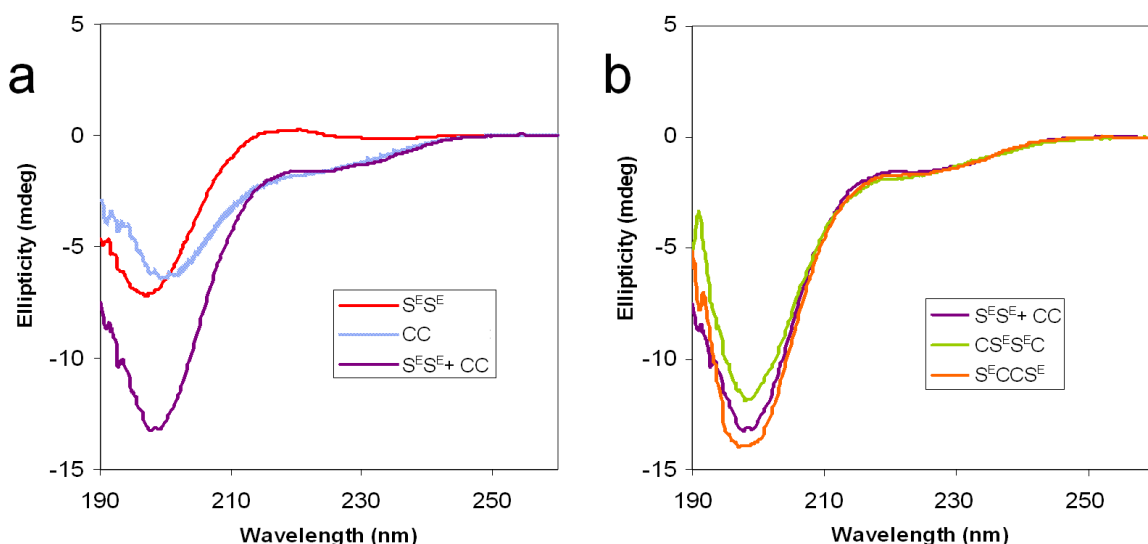


Figure 3.6 (a) The sum spectrum and separate spectra of 0.042 g/l isolated $S^E S^E$ block at pH 7 and of 0.058 g/l, isolated CC block at pH7. (b) The same sum spectrum compared to the spectra of $CS^E S^E C$ and $S^E CCS^E$.

The only difference between $CS^E S^E C$ and $S^E CCS^E$ seems to be the somewhat increased bundle formation upon drying of $S^E CCS^E$ fibrils. The dried fibrils of $S^E CCS^E$ (Fig. 3.3 d) frequently form bundles whereas bundles were seldom seen in dried $CS^E S^E C$ (Fig. 3.3 c). The bundles split and converge randomly (Fig. 3.3 d). Possibly, this could be due to a different behavior of CC blocks looping from S^E to S^E block in $S^E CCS^E$, or to the occurrence, now and then, of $S^E CCS^E$ molecules bridging two fibrils by insertion of their end blocks in different

fibrils rather than in the same fibril. By chemically attaching polydisperse polyethylene oxide (PEO) to the N- and C-terminal ends of (GAGAGAGE)₁₀ or (GAGAGAGE)₂₀, also the group of van Hest [26] could prevent isotropic aggregation of the silk-like polymer. This resulted in long fibrillar structures similar to ours, be it at high protein concentration and in crystallization-promoting solvent mixtures rather than water. It would be interesting to know whether such silk-PEO copolymers with an inverted S-PEO-S triblock order would behave similarly.

3.3.2 Effect of pH on secondary structure of CS^ES^EC and S^ECCS^E

Secondary structure changes induced by shifting the pH from 7 to 2 were followed by circular dichroism (CD). In agreement with the above-mentioned observation that fibrils were not seen immediately after acidification, the development of the spectra associated with fibrils shown in **Figure 3.5** took several hours. For comparison, two separate polymers, with a length and amino acid composition equal to the CC [11] or S^ES^E [12] blocks, were subjected to the same pH shift. At pH 7 or above, the CD spectra of CS^ES^EC, S^ECCS^E (**Fig. 3.5**) and the isolated S^ES^E (**Fig. 3.6 a**) [12] all revealed a combination of random and extended structure.

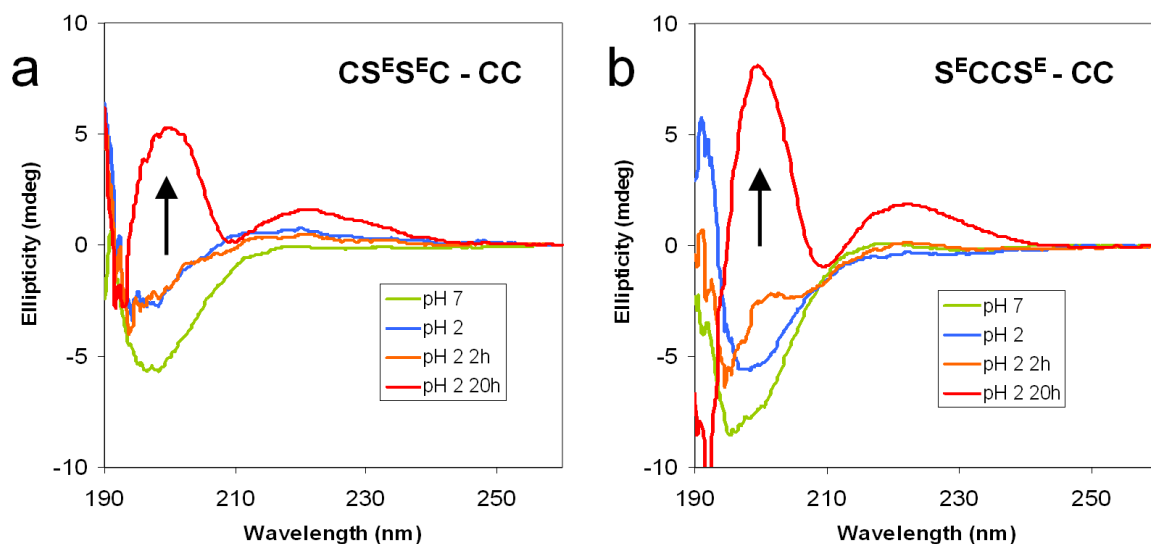


Figure 3.7 (a) The spectrum of 0.042 g/l central S^ES^E block at pH 2, by subtracting the spectrum of 0.058 g/l CC from the 0.1 g/l CS^ES^EC spectrum (b) Similarly, the spectrum of 0.042 g/l flanking S^E.

While S^ES^E became turbid due to random aggregation at low pH, both CS^ES^EC and S^ECCS^E formed a transparent gel or dispersion of fibrils, allowing the recording of UV-CD spectra down to at least pH 2 (**Fig. 3.5**). The secondary structure of the S^E block in CS^ES^EC

and S^ECCS^E can probably be followed at low pH by subtracting the CC spectrum from the CS^ES^EC and S^ECCS^E spectra. This can be done because the isolated CC molecule [11] retains exactly the same (unordered structure and) CD spectrum as depicted in **Figure 3.6 a** at low as well as at neutral to alkaline pH, and also because, at least at pH 7, the sum of the CC and S^ES^E spectra closely matched the CS^ES^EC and S^ECCS^E spectra (**Fig. 3.6 b**).

The result of such an S^E block-revealing subtraction is shown in **Figure 3.7**. The combination of a negative band around 208 nm and positive bands around 200 and 225 nm, seen at low pH in both the original and subtracted spectra, indicates the S^E block is rich in β -turns [28]. The CD spectra closely resemble those of isolated S^ES^E block dried from an aqueous medium onto a quartz substrate [12], and that of poly-(GAGAGAGX) with Tyr, Glu, His and Lys residues in consecutive X positions (YEHK) [27, 29], but differ from the CD spectrum of (GAGAGAGE)_n ribbons that were crystallized from a mix of methanol and formic acid [26]. YEHK [27, 29] was thought to form antiparallel β -sheets like the ones in methanol-formic-acid-derived (GAGAGAGE)_n crystals [26, 30], but it had an unusual CD spectrum, resembling ours (**Fig. 3.5**). The difference between the CD spectra of YEHK and normal β -sheet CD spectra was attributed to π -stacked arrays [29], but we find a very similar, unusual CD spectrum with only glutamic acid residues on the X position of the (GAGAGAGX)₄₈ repeat, which are not aromatic and do not possess such π -stacking capability.

We suggest that, apart from random structures, the (GAGAGAGX)_n repeats can adopt at least two different fibril forming conformations, depending not so much on the amino acid(s) on the X position [31], but more on the sample history. The first is an antiparallel β -sheet [26, 30], formed when crystallized from a methanol-formic-acid mixture, and the other is an illusive structure that according to CD spectra seems to be common for our (GAGAGAGE)₄₈ repeat and YEHK [29], both obtained from aqueous solution.

The CD spectrum obtained for the S^E block after self-assembly probably has a lower molar ellipticity than may be expected for non aggregated molecules with the same structure. This is because of the ‘spectral flattening’ effect [32]: if dimers or multimers occur (self-assembled fibrils are multimers) the molecules closest to the light source will absorb nearly all the light and cast a molecular shadow on the molecules that are more distant from the light source but within the same multimer. Thus, the effective molar ellipticity of the multimer is significantly lower than the molar ellipticity of the monomolecularly dissolved species [32].

We assume that no significant spectral flattening occurs for the C block spectrum upon self-assembly of the S^E block. Because of its hydrophilic nature under all conditions it will form a dilute corona with a much lower density than that of the tightly packed S^E block. Therefore we can subtract the C spectrum from the spectrum of the total block copolymer as described above.

3.3.3 Stacking and aggregation of an S^E block β -roll conformation according to molecular dynamic (MD) simulation

From CD spectroscopy (**Fig. 3.5**), we learned that the formation of the supramolecular ribbons observed in the micrographs (**Fig. 3.3**) is driven by the conformational change of the S^E block, from a random structure to a conformation that is rich in β -turns. However, because the CD spectrum of our ribbons did not match the CD spectrum of any known structure, the nature of this β -turn rich structure was still unclear. Therefore we employed molecular dynamic (MD) simulation to find possible S^E block structures (**Fig. 3.8, 9**) that are rich in β -turns and are likely to form ribbons with dimensions comparable to the ones we found in the micrographs (**Fig. 3.3, 4**).

First we simulated (GAGAGAGE)_n in the β -sheet structure that was reported for such a repeat, crystallized from organic solvents [30]. In water this β -sheet unfolded immediately. This was obviously not a structure that would be stable in aqueous environment. Then a bottom up approach was tried to reach a regular, and stable structure, rich in β -turns. The starting structures used for simulations were hairpins, with ψ and ϕ angles of the β -strands based on 1SLK or 2SLK structures [22] and a glutamic acid containing β -turn [23]. When simulations were run for 10 ns, the 1SLK structure was unstable, but the 2SLK structure adjusted itself by twisting, resulting in a stable, twisted hairpin (**Fig. 3.8 a**) with a turn resembling that of a previously published β -roll in 1KAP [33]. Using the 1KAP entries in the Protein Data Bank, a β -roll (**Fig. 3.8 b**) was created that appeared to be stable in simulation.

The characteristic features of β -rolls like ours (**Fig. 3.8 b**) are the two parallel β -sheets, oriented in opposite directions, through which the backbone of the polymer spirals forward, like a string wrapped at an angle over a ruler, with one β -strand in the one β -sheet and the next β -strand in the other, connecting the two oppositely directed β -sheets to each other on both sides, with a set of turns. As a result, hydrogen bonding connects any strand (i) not to the next strand (i+1) as in an antiparallel β -sheet, but to one strand further in the sequence (i+2).

Simulating a Stack of two of these β -rolls, led to two even more stable and regular β -rolls (**Fig. 3.8 c**). The turns of these β -rolls, with average ψ and ϕ angles of C α 2 (109.2, 111.9) and C α 3 (-132.2, 33.0), do not fall into any category of β -turn, and may account for the unusual CD spectrum.

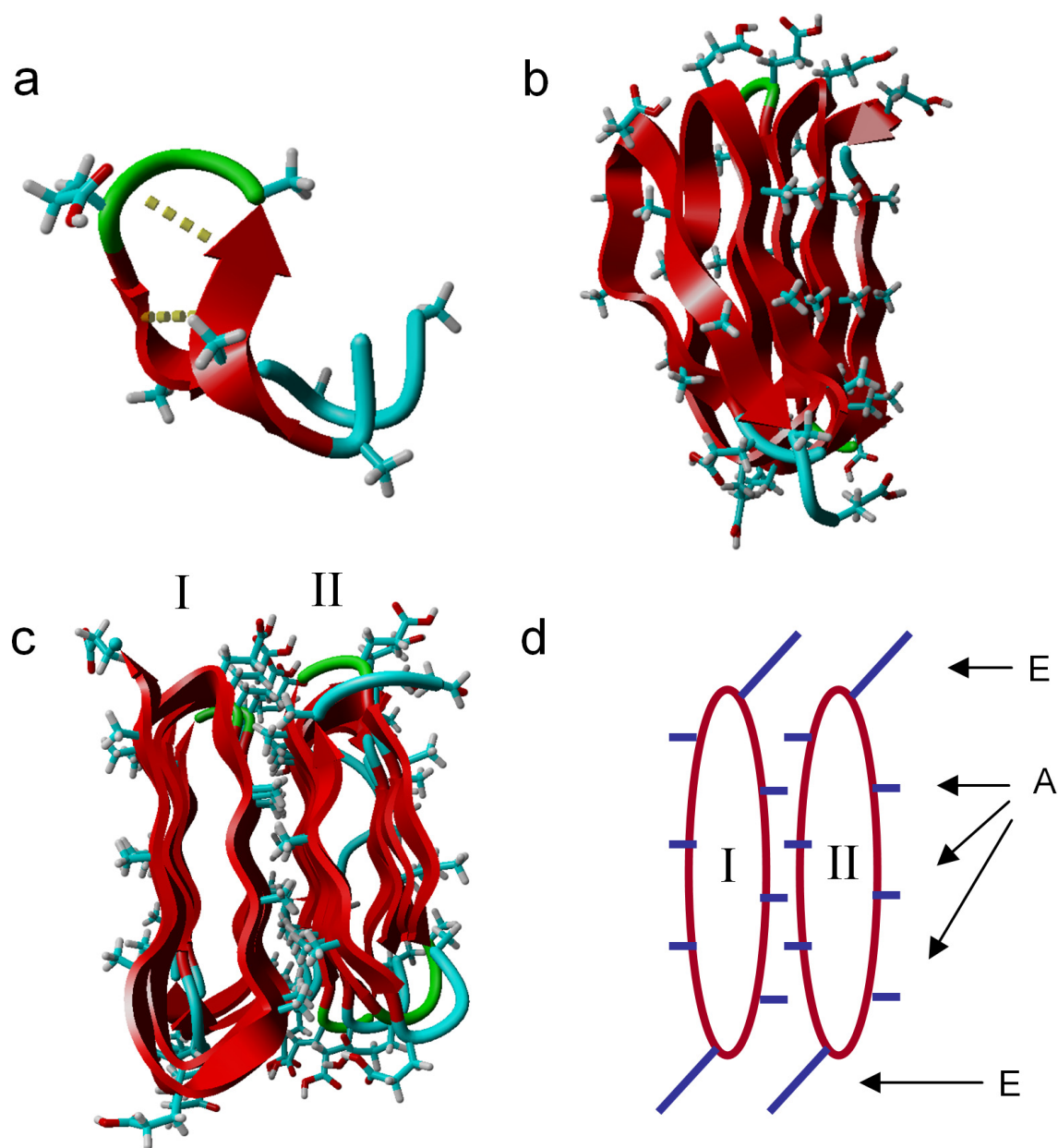


Figure 3.8 Stable structures of (GAGAGAGE) repeats resulting from MD simulation. (a) Stable twisted hairpin of the sequence GAGAGAGEGAGAGA with the glutamic acid in the turn and with hydrogen bond shown as dotted lines between the antiparallel β -sheets shown as ribbons. (b) Stable β -roll with turns like in figure 3.8a. (c) Two β -rolls (I and II) as in (b), stacked, stabilizing each others structure (d). Schematic representation of β -rolls (I and II) from (c), with the amino acid residue side chains of alanine (A) and glutamic acid (E) pointed out with arrows, to illustrate clearly how the alanine side chains (methyl groups) of the two β -rolls fit between each other.

For clarity, β -roll I was taken from **Figure 3.8 c** (omitting β -roll II) and tilted slightly (**Fig. 3.9**), to clearly show the regular spacing of the methyl (A) and the carboxylic acid (E) side groups of the alanine (A) and the glutamic acid (E) residues respectively. As depicted in the cartoon (**Fig. 3.8 d**) the regularly spaced methyl groups on both β -rolls fit neatly between each other, creating a nonpolar environment between the β -rolls. Many β -rolls may stack in the same way and form a hydrophobic S^E block ribbon with a hydrophilic, glutamic acid coated surface. The propagation direction of such a ribbon in **Figure 3.8 c** would then be from the left to the right. A ribbon based on these β -rolls does not have a front or back end. Both ends are equal. However, the edges of the ribbon differ in that all β -rolls start at the one edge and end at the other. They are all oriented in parallel, perpendicularly to the long axis of the ribbon. For example, both β -rolls in **Figure 3.8 c** have their N-terminus away from the viewer and have their C-terminus closest to the viewer. If they were oriented antiparallely, with the amino acid number increasing in opposite directions, the glutamic acid side chains would interfere with each other, as can easily be imagined from the cartoon (**Fig. 3.8 d**) by flipping β -roll II over a vertical axis.

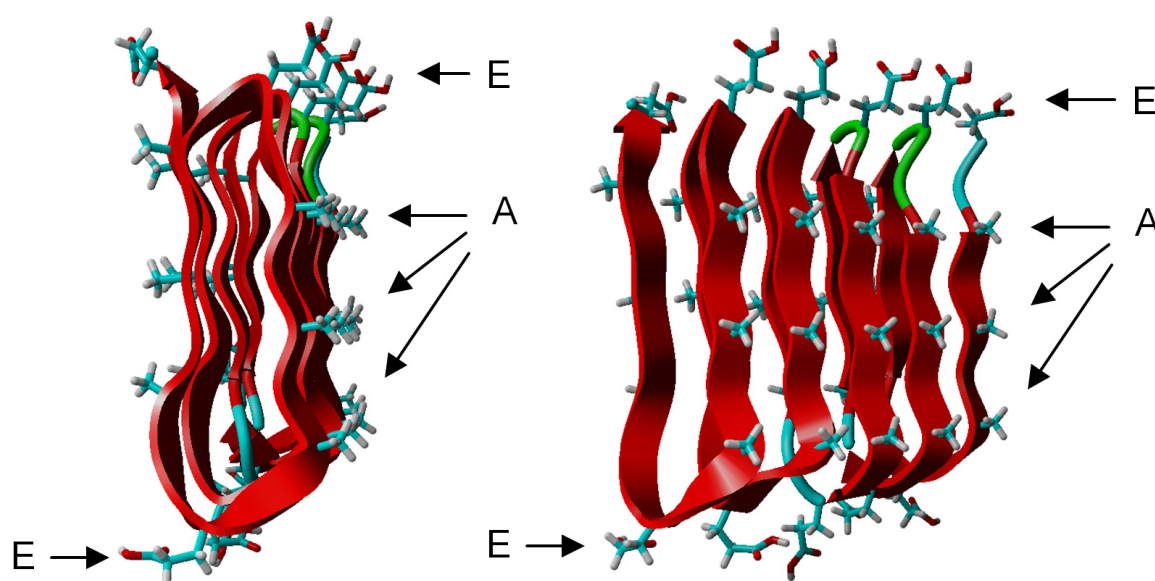


Figure 3.9 Different views of only β -roll I from figure 3.9 c, omitting β -roll II to have a clear view from different perspectives of the regular spacing of the methyl groups (A) on the flanks of the β -roll, and the glutamic acid side groups (E) located in the turns that link the β -strands of the one sheet with the other.

This water derived β -roll is a structure very different from the methanol-formic-acid-derived, antiparallel β -sheet [26, 30]. It is possible that, in water, YEHK [29] forms β -rolls similar to ours, accounting for the similar, unusual CD spectrum. However, the reported YEHK CD spectrum [29] is lower between 190 and 210 nm relative to the second maximum at 220 nm than our subtracted CD spectra (**Fig. 3.7**). This difference could have several explanations. One explanation might be the π -stacking of aromatic amino acid residues that are present in YEHK [29] but not in our S^E block. Another explanation might be that not all YEHK turns [29] are the same as in our β -roll, and depend on the amino acid residue in the turn. The YEHK CD spectrum [29] more closely resembles our non-subtracted spectra, suggesting that the YEHK CD spectrum [29] arose from a combination of β -rolls and random structures.

3.3.4 Pure CC block, a water swollen globule

For an accurate representation of the ribbons, it would be useful to have an estimate of the size and water content of the hydrophilic C block ribbon corona. Using dynamic light scattering (DLS), the hydrodynamic radius of dissolved CC molecules at pH 2 was determined. Analyzing the DLS data with the cumulant method resulted in an average first, second and third order cumulant of 7.2, 6.1 and 5.7 nm respectively. The difference between these cumulants indicates some polydispersity. In this case, taking the first cumulant of the second order cumulant fit (6.1 nm) would lead to an overestimation of the CC molecule radius. The contin method gives a more accurate radius than the cumulant fit, because contin uses populations of different particle sizes to fit the DLS data. According to the contin fit, the hydrodynamic radius of the CC molecule was 5.47 nm. The size of the C block is consistent with a swollen hydrophilic random coil containing 94% water.

3.3.5 Tentative model of CS^ES^EC ribbon

A representation of a CS^ES^EC ribbon (**Fig. 3.10**) was made based on the results mentioned above. In **Figure 3.10** such a ribbon is depicted with ribbon height (h) 2.7 nm and ribbon width (w) 11.6 nm. On the front a β -roll (green, striped) is shown, with two blue, dangling, hydrophilic C blocks. The β -roll is stacked onto a large number of other β -rolls, forming the partially hydrophobic ribbon core depicted as a green bar inside a hydrophilic (blue) corona of C blocks.

To calculate the expected core dimensions, we used the MD results of the two stacked β -rolls (**Fig. 3.8 c**). These β -rolls have an average height of 2.7 nm, an average interstrand distance of 0.48 nm and one β -roll repeating in the fibril direction every 0.95 nm. This is the same for any size S^E -type β -roll. An $S^E S^E$ middle block with 48 repeats, and an interstrand distance of 0.48 nm will result in a β -roll length, and therefore, also a ribbon width (w) of 11.6 nm. For a single S^E block as in $S^E C C S^E$, we expect β -rolls of half that size.

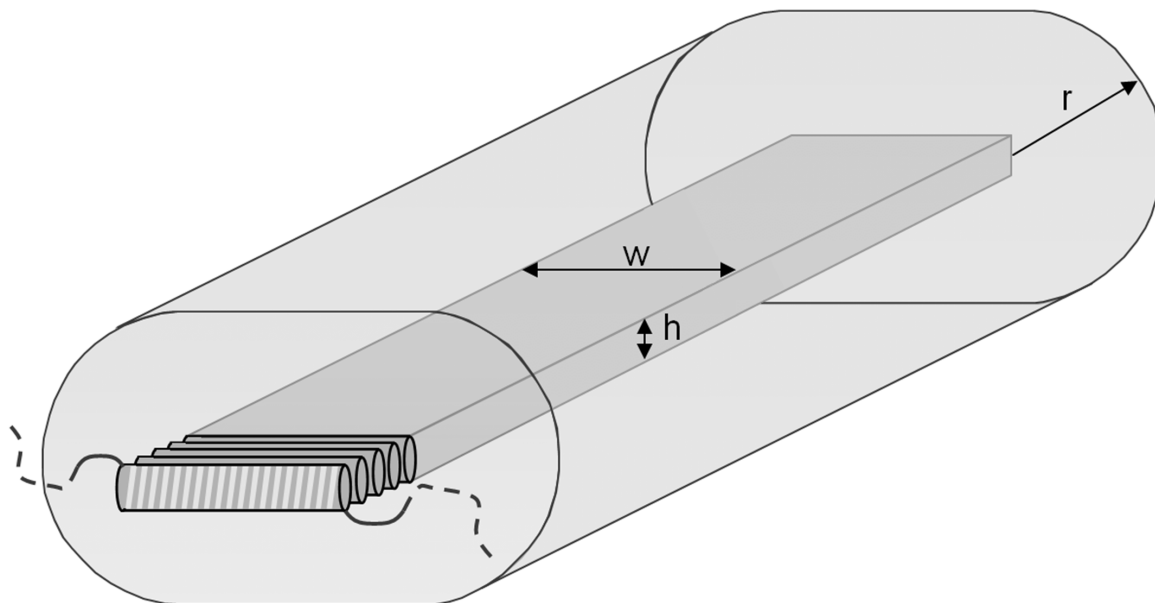


Figure 3.10 Tentative structure of the $C S^E S^E C$ fibrils formed at low pH. A compact stack of $S^E S^E$ block β -rolls (on the front), forms the ribbon shaped core. The β -rolls repeat every 0.95 nm of the ribbon. The ribbon has a height (h) of 2.7 nm and width (w) of 11.6 nm. These dimensions were extrapolated from MD results. The radius (r) of the hydrophilic C block corona is approximately 12 nm and is defined as the distance to the central plane in the core. The separate CC block contains around 94 % water and 6 % protein based on DLS measurements.

The corona dimensions were estimated assuming that the C blocks surrounding the core have the same density as the separate CC blocks measured with DLS. The globular volume, $4/3\pi * 5.47^3$, of the CC block could be converted to the volume of a disc with a thickness, 0.95 nm, equal to the repeat distance of the β -rolls. The surface of this disc is, according to **Figure 3.10**, equal to the summation of the surface of a circle with radius r , and a rectangle with a width, w , and height, $2r-h$. So, solving the equation of the two differently formed volumes: $(\pi r^2 + (2r-2.7) * 11.6) * 0.95 = 4/3\pi * 5.47^3$ gives a radius r as defined in **Figure 3.10** of 12.2 nm. Obviously, one C chain may cross over into another imaginary disc, but the estimate holds if the density of the C blocks in the corona is the same as that for a single globular CC

molecule. However, because the C blocks are packed closely together, the corona might be slightly denser than, and may have a different form from our representation.

If the two S^E blocks of one $S^E CCS^E$ molecule would aggregate freely with S^E blocks of other $S^E CCS^E$ molecules, one would expect fibrils or even molecular networks. Such networks were not observed, but only linear fibrils. It might be so, that the two S^E blocks in $S^E CCS^E$, separated by the central, flexible CC block, first aggregated with each other, possibly stacking as two β -rolls, before being integrated into a fibril. Considering that one S^E block β -roll has half the length of a central $S^E S^E$ block β -roll, this would lead to ribbons with half the width of the $CS^E S^E C$ ribbons. No convincing evidence for such a difference in width could be gathered from micrographs, so small angle x-ray scattering (SAXS) was used to determine the sizes of the ribbons.

3.3.6 Ribbon dimensions of $CS^E S^E C$ and $S^E CCS^E$ measured with SAXS

SAXS measurements were performed on 20 g/l gelled $CS^E S^E C$, and on 20 g/l $S^E CCS^E$, both at pH 1.5 (**Fig. 3.11**). Intensity as a function of q was determined between 0.12 and 1.23 nm^{-1} (**Fig. 3.11**). The slope between 0.5 nm^{-1} and 1 nm^{-1} was -2 , meaning that the measured objects were solid flattened particles. In the micrographs the objects appeared to be several μm long. Therefore we assumed that the objects were ribbons.

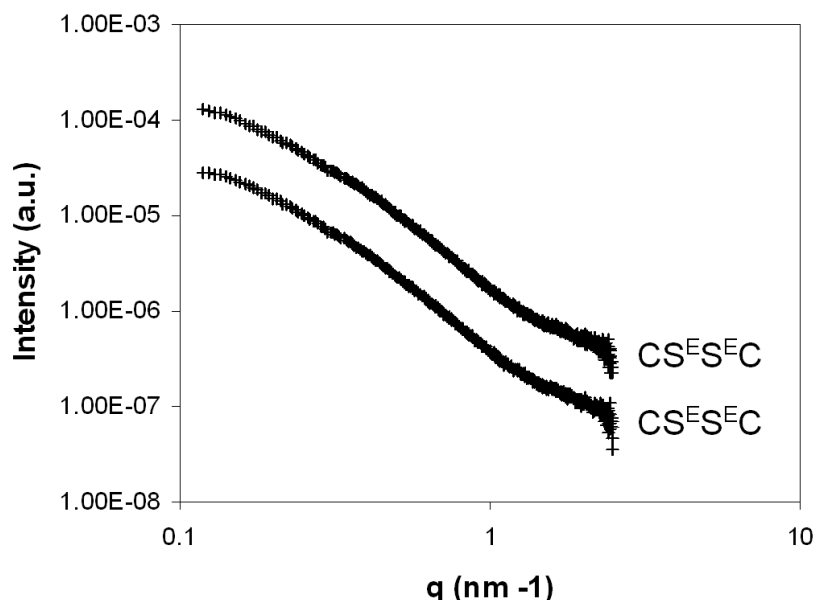


Figure 3.11 SAXS measurements of 20 g/l $CS^E S^E C$ and 20 g/l $S^E CCS^E$ hydrogels, both at pH 1.5, showing Intensity in arbitrary units as a function of q . The curve of $CS^E S^E C$ was shifted downward for clarity of the illustration. Guinier analysis revealed ribbon shaped objects with dimensions as defined in figure 3.6: a height (h) of 2.8 nm and a width (W) of 13.6 nm for $CS^E S^E C$ and an height (h) of 2.8 nm and a width (w) of 14.3 nm for $S^E CCS^E$.

Guinier analysis revealed that $CS^E S^E C$ formed ribbons with height (h) and width (w) as defined in **Figure 3.10** of 2.8 nm and 13.6 nm respectively, which is in good accordance with the 2.7 and 11.6 nm that we estimated for the core. According to estimates from DLS, the isolated C block is very dilute (containing 94% of water molecules). Assuming that in the fibril the C block is similarly dilute, the main electron density difference is at the surface of the core, rather than of the corona. So, it is correct to assume that the sizes measured with SAXS are mainly those of the core. Still the C block could contribute slightly to the measured sizes, especially if the material distribution in the corona is asymmetrical, being denser at the core and more dilute further out from the core. This can account for the small differences between the sizes, estimated from MD, and the (larger) sizes measured with SAXS. We expected $S^E CCS^E$ ribbons to have a height equal to, and a width approximately half of that of $CS^E S^E C$ ribbons. After Guinier analysis of the SAXS data of $S^E CCS^E$, we did indeed find a height of 2.8 nm but to our surprise, the ribbon width was 14.3 nm, which is more or less the same, actually even slightly larger than that of $CS^E S^E C$ ribbons. If our β -roll model is correct, this would imply that the width of the ribbon is still equal to the length of two S^E blocks.

3.3.7 Tentative model of $S^E CCS^E$ ribbon based on DLS, MD and SAXS

We propose a structure (**Fig. 3.12**), based on β -rolls, for $S^E CCS^E$ forming ribbons with dimensions consistent with SAXS measurements. In this zipper-like structure, the core is made up of two parallel stacks of β -rolls. Of all molecules, the C-terminal end S^E blocks are located in one stack, while the N-terminal S^E blocks are in the other stack. Because of the directionality of the β -rolls, C-terminal and N-terminal β -rolls pair up to form one discontinuous β -roll, the C-terminus being connected to the N-terminus by hydrogen bonding in a head to tail fashion. Hydrogen bonding probably occurs between the first two and the last two octapeptide repeats of the molecules. Thus, the two stacks in the core are closely juxtaposed and zipped together by N-C-terminal H-bonding. Because the middle CC block connects both of the S^E end blocks, the CC block forms a loop connecting the outer edges of the zipper-like ribbon (**Fig. 3.12**). The CC middle block also prevents further head to tail assembly of the β -rolls, perpendicular to the long axis of the ribbon. This delimits the sides of the ribbons and prevents two adjacent ribbons from zipping together along their small edges.

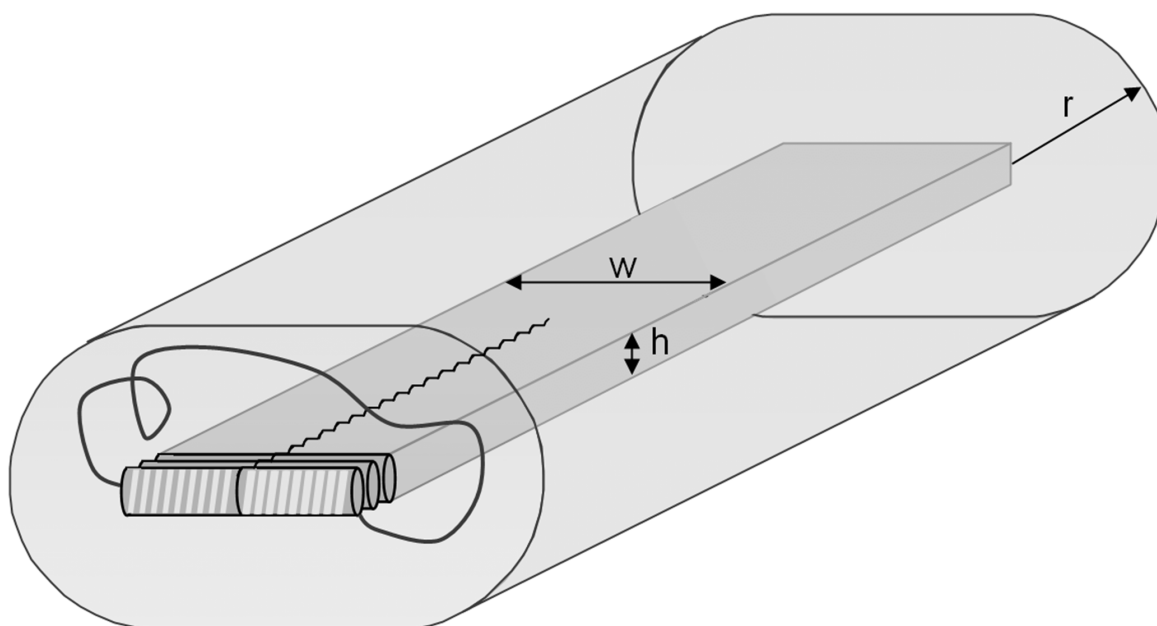


Figure 3.12 Tentative (zipper-like) structure of the $S^E CCS^E$ fibrils formed at low pH. Upon acidification, the two separate S^E blocks contained in one $S^E CCS^E$ molecule self-assemble on the end of the compact, ribbon shaped core of β -rolls. All β -rolls are oriented in parallel with the CC hydrophilic middle block, connecting the two S^E blocks located in the core. A “CC loop” must then be connected to both edges of the ribbon. Analogous to figure 3.6, the ribbon should have a height (h) of 2.7 nm and width (w) of 11.6 nm. The radius (r) of the hydrophilic C block corona is 12 nm and is defined as the distance to the central plane in the core. The free CC block contains around 94 % water and 6 % protein according to DLS measurements.

3.4 Conclusion Section I

3.4.1.1.1 pH induced self-assembly of the negatively charged $CS^E S^E C$ and $S^E CCS^E$

The produced protein block copolymers do not behave like traditional block copolymers in the sense that they do not form globular micelles, flower-like micelles or molecular networks. Instead they form the well-defined ribbons observed in the micrographs (**Fig. 3.3**). This difference is due to the S^E block that self-assembles into the defined ribbon-like core, opposed to polymers that aggregate amorously into a globular micellar core. The remaining similarity to traditional block copolymers is that the $CS^E S^E C$ ribbon has a hydrophobic S^E block core and a hydrophilic C block corona. Similarly $S^E CCS^E$ ribbons are covered with hydrophilic loops just like BAB block copolymer flower-like micelles.

The conformational changes of the S^E block during aggregation into well defined ribbons, could be followed with CD spectrometry (**Fig 3.5, 6, 7**). CD spectrometry revealed that S^E is indeed the aggregating block, and also that the C block retains its random coil conformation under all conditions tested. Although swelling or collapsing of a random coil can not be detected with CD spectrometry, the C block must remain hydrophilic under all tested conditions, otherwise precipitation would have been observed. Still, the unusual CD spectrum that we obtained for the S^E block in its ribbon form was different from the CD spectrum of similar ribbons [26] reported earlier. It seems that (not regarding random structures) there are at least two possible conformations for (GAGAGAGX)_n repeats, depending not so much on the amino acid(s) on the X position [31], but more on the sample history. From a methanol, formic acid mixture, an antiparallel β -sheet structure is obtained [26, 30], but in aqueous solutions, a different structure is obtained. The nature of this structure was to date not resolved [27, 34]. Therefore molecular dynamic (MD) simulation was employed to find structures that are likely to form ribbons that are stable in water with dimensions comparable to the ones we found in the micrographs (**Fig. 3.3**).

With MD simulations, only one stable structure for the uncharged S^E block in aqueous environment was found: a β -roll (**Fig. 3.8, 9**), with an unusual turn, that could account for the obtained unusual CD spectrum. This β -roll might also be the structure of the poly-YEHK molecule [27, 29], for which a CD spectrum [29] similar to the $S^E S^E$ block spectrum was observed. In simulation, stacking the β -rolls, stabilized their structure even further (**Fig. 3.8 c**). Using the dimensions of the β -rolls in **Figure 3.8 c**, that contain only 10 (GAGAGAGE) repeats, the length, height, and repeat length in a stack of β -rolls, of the 48 repeat long $S^E S^E$ block were estimated to be 2.7 nm, 11.6 nm, and 0.95 nm respectively. Consequently, the cross-sectional dimensions of a stack of $S^E S^E$ block β -rolls were estimated to be 2.7 nm, and 11.6 nm respectively. Estimates for the hydrophilic corona around the ribbon were derived from DLS measurements of the pure CC molecule. Finally a tentative model could be constructed of the $CS^E S^E C$ ribbons (**Fig. 3.10**), containing an $S^E S^E$ block core of stacked β -rolls and a hydrophilic corona of C blocks. SAXS results confirmed the size of the ribbon core, with a ribbon height and width of 2.8 nm and 13.6 nm respectively. The measured core size is expected to be slightly larger than the estimated core size, because of a small contribution of the corona to the SAXS measurement. Conversely, AFM measurements gave

a fibril thickness of 1.5-2.0 nm. This might be due to conformational changes in the fibril resulting from drying, or adsorption onto the silica surface of the sample.

S^ECCS^E was expected to have a core of half the width of the CS^ES^EC core because the separated S^E blocks of S^ECCS^E have half the size of the central S^ES^E block in CS^ES^EC . To our surprise we found that the dimensions of the core for both S^ECCS^E and CS^ES^EC ribbons are almost the same. We proposed a zipper-like structure (**Fig. 3.12**) in which the β -rolls connect in the middle of the ribbon in a head to tail fashion, explaining the width of the ribbon. The hydrophilic loops of central CC blocks prevent further aggregation in the lateral direction of the ribbon. So, it appears that the protein block copolymers described in this work, in contrast to traditional block copolymers, have the tendency to form a new class of structures: highly defined self-assembled ribbons with a crystalline core, instead of globules with an amorphous core.

3.5 Results and discussion Section II

pH induced self-assembly of S^H containing protein polymers

3.5.1 Supramolecular ribbons of CS^HS^HC and S^HCCS^H at high pH

The histidine containing protein block copolymers CS^HS^HC and S^HCCS^H were created similar to the negatively charged CS^ES^EC and S^ECCS^E but instead, positively charged, so that they would aggregate and form gels at high pH rather than at low pH. These positively charged molecules did indeed form gels at high pH (pH12), and nanoscopic fibrils (**Fig. 3.13, 14**) again caused gelling.

With AFM (**Fig. 3.13 a, b**), fibrils were observed with a height of 1.8-2.2 nm and a tip convoluted, hence apparent width of several nanometers, which was similar to the dimensions that we found for CS^ES^EC and S^ECCS^E (height 1.5-2.0 nm, width several nm), also resembling those fibrils for PEO-conjugated poly(GAGAGAGE) [26], crystallized from organic solvents, and poly-YEHK, deposited from water [27]. The height is less than one would expect for a β -roll as described above, but absorption to the silica surface and drying might have caused conformational changes, which can explain the discrepancy.

There were much less fibrils adsorbed to the silica surface than in the case of CS^ES^EC and S^ECCS^E . Several empty μm^2 had to be scanned before finding some fibrils. Possibly fibrils of CS^HS^HC or S^HCCS^H form slower than fibrils of CS^ES^EC or S^ECCS^E , leading to less

fibrils in solution at the time that the silica was dipped in the fibril suspension to take a sample. The presence of supramolecular fibrils was confirmed with transmission electron microscopy (TEM) of negatively stained preparations (**Fig. 3.13 c,d**). However, these images might not be representative of the fibrils in solution, because, apart from drying, the sample was exposed to a staining solution with uranyl acetate at pH 3.8. The low pH might have caused the fibrils to charge up and fall apart, or change their form (which would not happen for the negatively charged protein block copolymers at low pH). However, during the sample preparation, material was first adsorbed and then stained, so fibrils seen in the TEM images were actually present in solution. The large amount of randomly structured material observed might be fibrils that fell apart, or $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ adsorbed randomly from solution.

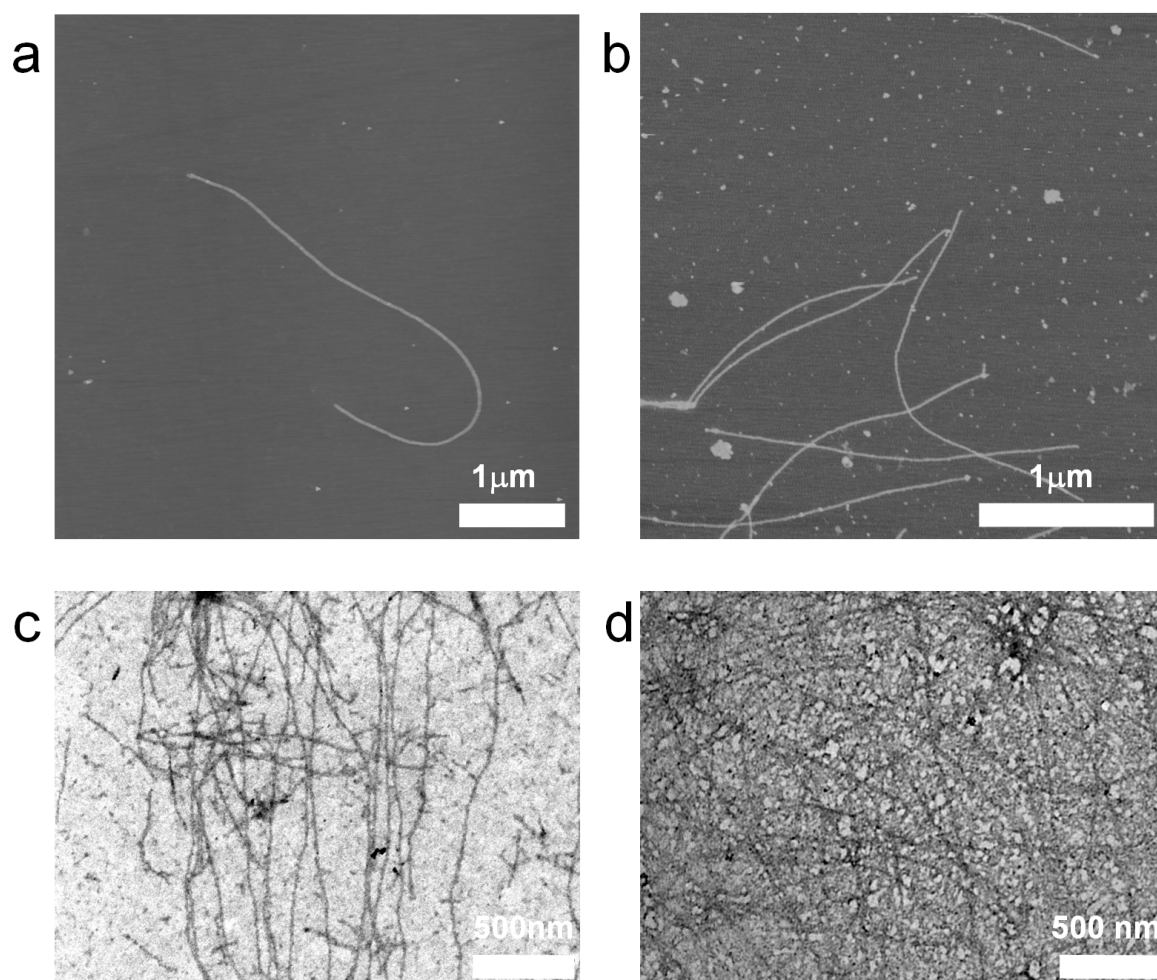


Figure 3.13 Micrographs of supramolecular $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ fibrils formed at pH 11. (**a,b**) Atomic force micrographs of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ fibrils, respectively, deposited on silica. Higher objects are lighter in color. (**c,d**) TEM images of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ fibrils, respectively, deposited on Formvar.

Further evidence for fibril formation in solution is given both by cryogenic (cryo)-TEM and by AFM micrographs. In the cryo-TEM micrograph (**Fig. 3.14**), the occurrence of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ fibrils in solutions that were frozen after overnight incubation at pH 12 reveals that the formation of fibrils observed in TEM images (**Fig. 3.13 c,d**) was not merely due to the process of drying the alkaline protein polymer dispersions for subsequent TEM analysis. In cryo-TEM (**Fig. 3.14**), the fibrils were clearly visible, but, due to lack of contrast, not visible enough for conclusions to be drawn as to their exact size and morphology. In AFM images (**Fig. 3.13 a,b**) the apparent increase in fibril thickness at intersections of fibrils is also in support of the fibrils forming in solution before being deposited on the silica.

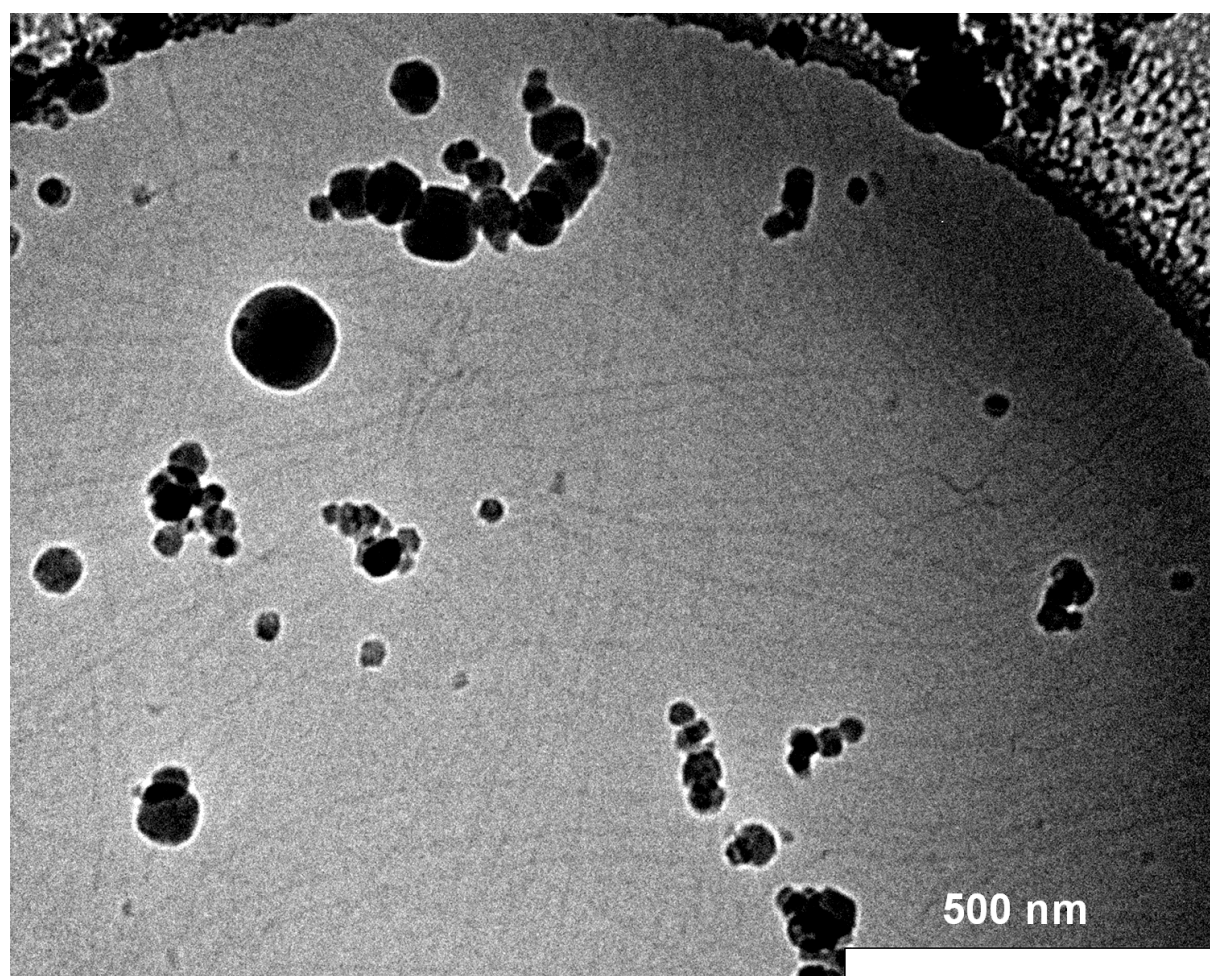


Figure 3.14 Cryo-TEM micrograph of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ fibrils in copper grid (The large, dark objects are frozen condensate). The fibrils were deposited from a 0.1 g/l protein solution.

3.5.2 Effect of pH on secondary structure of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$

The molecules $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ presented an opportunity to study the influence of amino acid residue on the conformation of the S block. The intention is to study whether at high pH, the uncharged S^{H} block forms a β -roll similar to that of the S^{E} block at low pH. This may be so, if the uncharged amino acid residue in the turn has little influence on the conformation of the block.

Like for the S^{E} containing products, the CD spectra of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ are assumed to be a proportionate sum of always the same CC block spectrum and a variable $\text{S}^{\text{H}}\text{S}^{\text{H}}$ block spectrum which depends on conditions. We also assume that spectral flattening [32] may occur for the S^{H} blocks but not for the C blocks.

CD spectrometry was used to monitor the conformational response of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ to high pH. At neutral pH, when the S^{H} blocks are positively charged, both S^{H} containing molecules displayed a CD spectrum (**Fig. 3.15 a,b**) of random and extended structures, similar to the CD spectra of the negatively charged S^{E} containing molecules (**Fig. 3.5**). In both cases charge repulsion could cause the random extended structures of the S blocks. After increasing the pH to pH 12 and waiting for 2 hours, both $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ displayed a CD spectrum with a maximum at 220 nm, a minimum at 205 nm, and possibly a maximum at 198 nm. Generally, one could say that, below 220 nm, the S^{H} CD spectrum is similar in form to that of S^{E} (**Fig. 3.5**), but has a more negative ellipticity. This could be due to slow conformational changes that had not yet been completed when the CD spectrum was taken. Completing these changes might finally led to the same CD spectra as in **Figure 3.5**. However, after waiting for an additional 1.8 hours, exactly the same CD spectra were obtained as before, indicating that conformational change had ceased.

After subtraction of the CC block spectrum, (**Fig. 3.15 c,d**) the S^{H} spectrum was similar to that of S^{E} and to that reported for YEHK [29], but more similar to the YEHK spectrum than to the S^{E} spectrum. π -stacking of the aromatic side groups of both S^{H} and YEHK [29] could cause their CD spectrum to deviate from the S^{E} CD spectrum. S^{H} and YEHK [29] could also have a conformation somewhat different than S^{E} e.g. a mix of β -rolls and more unordered structures.

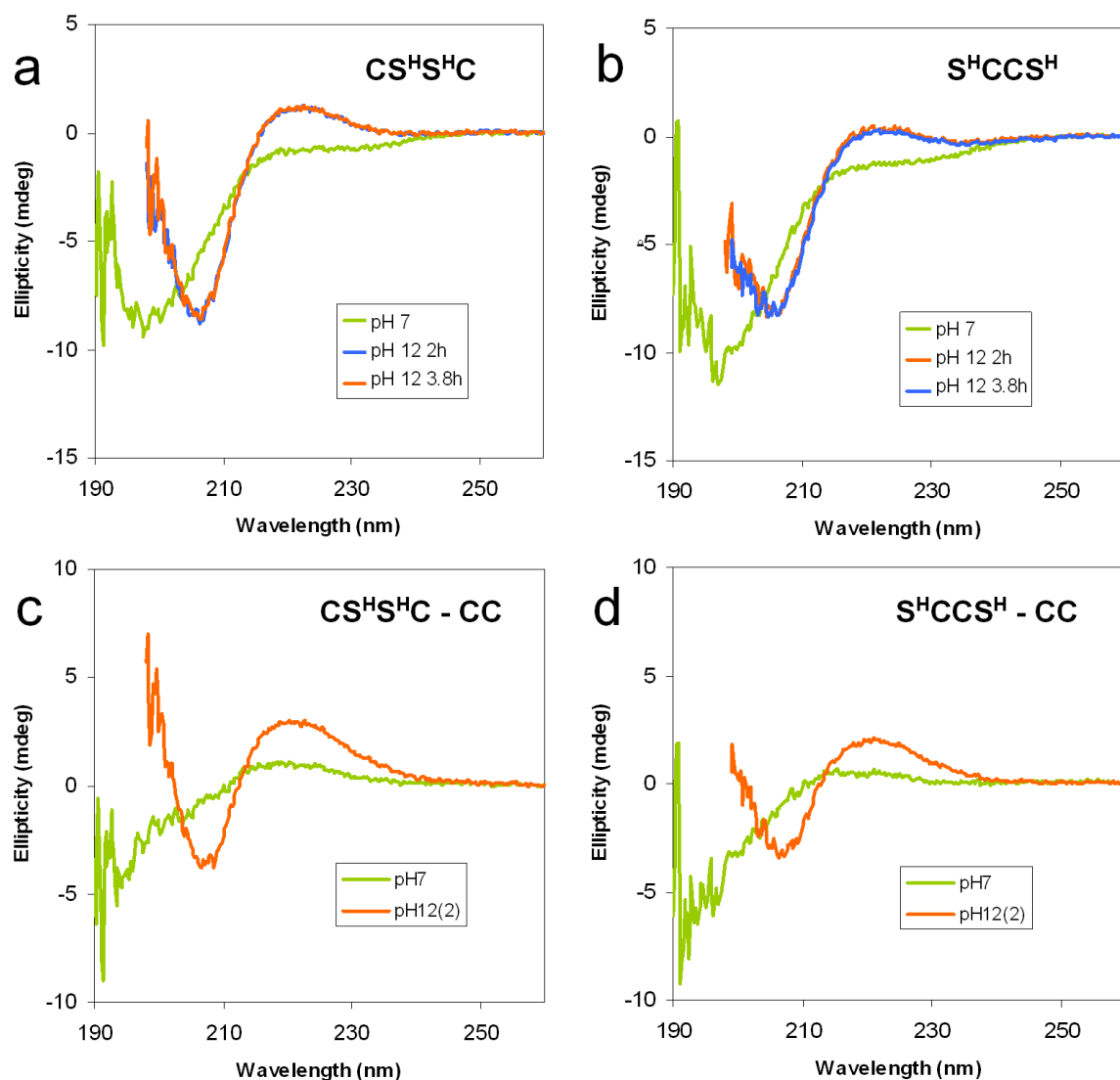


Figure 3.15 CD spectral changes reflecting structural changes induced by shifting the pH from 7 to 12 of: (a) 0.1 g/l $CS^H S^H C$ and (b) 0.1 g/l $S^H CCS^H$. (c) Analogous to figure 3.4, the spectrum of 0.042 g/l central $S^H S^H$ block by subtracting the 0.058 g/l CC spectrum from the 0.1 g/l $CS^H S^H C$ spectrum (d) Similarly, the spectrum of 0.042 g/l flanking S^H blocks.

3.5.3 SAXS reveals a mixture of ribbons and globules for $CS^H S^H C$ and $S^H CCS^H$

SAXS measurements were taken of 19.4 g/l $CS^H S^H C$, and of 19.4 g/l $S^H CCS^H$, gels pH12 (**Fig. 3.16**). The obtained data could be interpreted in two different ways. The gel either consisted of nanoscopic rods, or it consisted of a mixture of nanoscopic ribbons and globules. Both would yield the same SAXS spectrum.

According to the first interpretation, Guinier analysis revealed that both $CS^H S^H C$ and $S^H CCS^H$ gels consisted of rod-like filaments with a diameter close to 3 nm, $CS^H S^H C$ rods

having a slightly smaller diameter than S^HCCS^H rods. The presence of filaments was confirmed with microscopy (**Fig. 3.13, 3.14**). However we could not confirm their shape, because (dry) AFM (**Fig. 3.13 a,b**) and (dry) TEM (**Fig. 3.13 c,d**) may not be representative of the structure in solution, and because the contrast of our filaments in cryo-TEM (**Fig. 3.14**) was not high enough to draw conclusions about the cross-sectional sizes and morphology of these filaments.

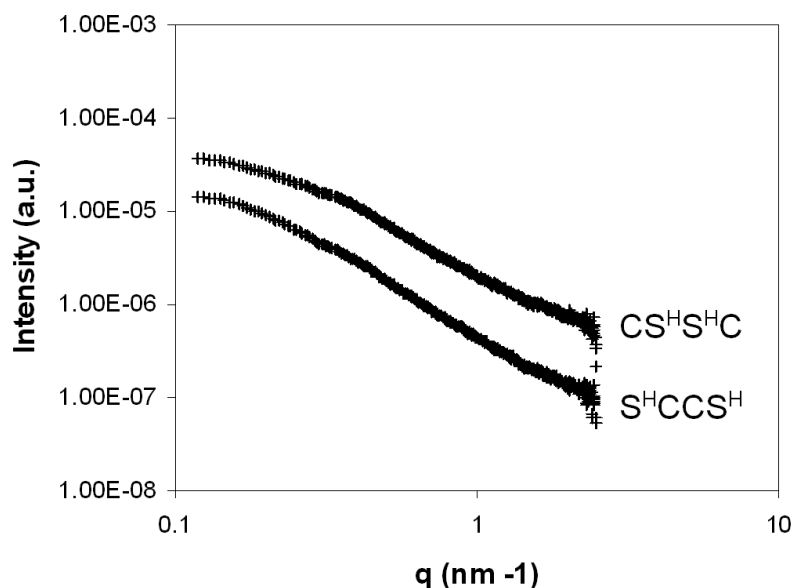


Figure 3.16 SAXS measurements of 19.4 g/l $CS^H S^H C$ and 19.4 g/l $S^H CCS^H$ hydrogels, both at pH 12, showing Intensity in arbitrary units as a function of q . The curve of $CS^H S^H C$ was shifted downward for clarity of the illustration. Guinier analysis revealed rod shaped objects with a radius of 3 nm. However, a combination of ribbons and globules could also be fitted to the data.

The glycine-alanine repeat is not expected to behave differently at high pH than at low pH, because it does not bear any charge that can respond to pH. Similarly, there is not much reason for the uncharged histidine at high pH to have a different conformational preference than the glutamic acid at low pH. Further more, the TEM micrographs (**Fig. 3.13 c,d**) revealed a mixture of absorbed filaments (possibly ribbons) and random aggregates. This could very well be the cause of a SAXS spectrum that can be fitted with a mixture of ribbons and globules. A mixture of β -rolls and unordered structures is also a possible explanation for the deviation of the S^H CD spectrum from that of S^E . Taking these considerations into account, the second interpretation of the SAXS spectrum: ribbons and random coils, seems the most likely. The reason why, for the S^H containing products, part of the molecules are kinetically trapped in random aggregates must be sought in the only difference with the S^E

containing products, which is the presence of histidine instead of glutamic acid in the silk-like block. In its uncharged state, the histidine residue might be more “sticky” because of hydrophobicity and π -stacking than the glutamic acid residue, leading to the kinetically trapped random structures.

3.6 Conclusion Section II

Having another amino acid residue on the X position of a (GAGAGAGX)_n repeat does have an influence on the formed structure. The polymers CS^HS^HC and S^HCCS^H at high pH possibly form combinations of ribbons and random structures. This is consistent with the CD, and SAXS measurements, and TEM micrographs. It is not inconsistent with the cryo-TEM observations, because with this technique only the dense fibril cores are observed and the random structures remain invisible. It is also not inconsistent with the AFM micrographs because fibrils and random aggregates may have a different affinity to the silica surface used for AFM, and the contact time with fibril containing solution was much shorter for the silica than for the Formvar used for TEM imaging. The ribbons are expected to have the same β -roll containing structure as the S^E blocks, but having a histidine instead of a glutamic acid seems to render the S^H block more “sticky” than the S^E block, leading to kinetically trapped random structures coexisting with the ribbons in the same sample.

The CD spectra of S^E, S^H and YEHK [29] resembled each other strongly, but the S^H spectrum showed the strongest resemblance to the YEHK spectrum [29]. Most likely the S^H block and the YEHK molecule [29] both form a combination of β -rolls, similar to the ones of S^E, and kinetically trapped random structures, since their CD spectra would come very close to a CD spectrum composed of a random, and (S^E-type) β -roll spectra.

3.7 Results and discussion Section III

Co-assembly of S^E and S^H block copolymers with polyelectrolytes or each other

3.7.1 Co-assembly of CS^ES^EC or CS^HS^HC with zwitterionic POWT

Until here, only self-assembly of the protein polymers at extreme pH has been discussed. Although, at pH 2 or at pH 12, there might be applications for such self-assembled networks, usefulness of these self-assembling molecules would be increased if self-assembly could be triggered in a large pH range. For self-assembly to occur, the charge of the self-

repelling S^E or S^H blocks must be neutralized. At a more moderate pH this could be done by adding an oppositely charged polymer (polyelectrolyte), leading to co-assembly of the polyelectrolyte with the charged S^E or S^H blocks, or by mixing $CS^E S^E C$ or $S^E CCS^E$ with $CS^H S^H C$ or $S^H CCS^H$.

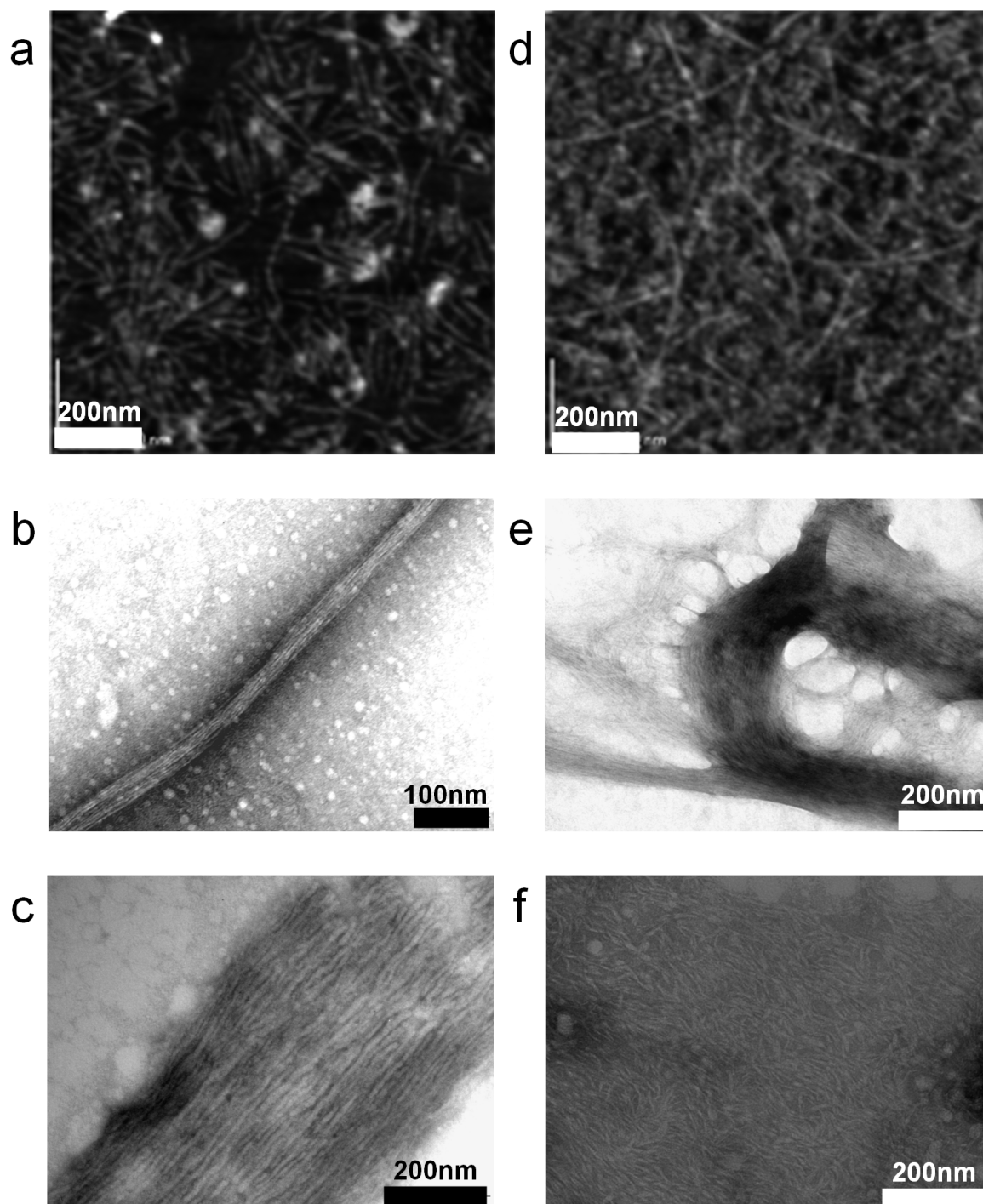


Figure 3.17 Micrographs of dried supramolecular $CS^E S^E C$ and $CS^H S^H C$ co-assemblies with POWT. (a and d) Atomic force micrographs of POWT containing $CS^E S^E C$ and $CS^H S^H C$ fibrils respectively, deposited on mica. (b, c) TEM images of $CS^E S^E C$ fibrils deposited on Formvar, from different locations

of the same grid. (e, f) TEM images of CS^HS^HC fibrils deposited on Formvar, from different locations of the same grid.

To compare the behavior of CS^ES^EC and CS^HS^HC, this was done with a water-soluble 15 monomers long polythiophene (POWT) (Fig. 1 a) in which every monomer has a zwitterionic substituent, that can be both positively and negatively charged. This property enables POWT to adjust its charge to that of both the negatively and the positively charged protein polymers. In both cases, co-assembling resulted in the observed fibrils (Fig. 3.17). In the co-assemblies, CS^ES^EC and CS^HS^HC, did not behave differently. Both protein-POWT co-assemblies had similar dimensions.

In the AFM micrographs (Fig. 3.17 a,d) POWT containing fibrils are found, both for CS^ES^EC (Fig. 3.17 a) and CS^HS^HC (Fig. 3.17 d). They were deposited on mica from a solution containing protein polymer and POWT in a charge ratio of protein/POWT = 1.5. TEM images (Fig. 3.17 b,c,e,f) reveal similar fibrils, deposited from the same solutions, which either dried in bundles (Fig. 3.17 b,c,e), or as dispersed nano-fibrils (Fig. 3.17 f) on the Formvar surface of the TEM grid.

The micrographs of the protein-POWT complex fibrils show larger amounts of fibril than the micrographs of the single protein components at extreme pH. This difference in fibril amount is due to a difference in sample preparation. Protein-POWT complex fibrils were deposited on the substrate by drying a drop of the suspension on the surface, while the single component fibrils were absorbed from solution after which the surface was rinsed. Additional AFM micrographs (data not shown) of the protein-POWT complex fibrils were prepared similarly to the ones in Figures 3.17 a and b, but instead from 10 times diluted dispersions. This yielded separate fibrils, enabling us to measure their dimensions. Both CS^ES^EC and CS^HS^HC fibrils with POWT had a tip convoluted, hence apparent width of 25-30 nm. The CS^ES^EC-POWT complex fibrils had a height of about 0.7 nm and the CS^HS^HC-POWT complex fibrils had a height of about 0.9 nm. These measured heights are less than the 1.5-2 nm measured for pure CS^ES^EC fibrils and the 1.8-2.2 nm measured for pure CS^HS^HC fibrils at extreme pH. The height of the POWT containing dry fibrils can not accommodate an S^E-type β -roll standing upright. The only way such a β -roll would fit is if it lay flat on the surface.

Another structure than a β -roll may also be possible. In aqueous solution, the structure for the POWT containing complexes is still unresolved. Structural indications could not be obtained with CD spectrometry, because the polythiophene disturbed the CD measurements. Also no cryo-TEM or SAXS measurements have been performed yet. However, POWTs

fluorescence emission spectrum (data not shown) did confirm co-assembly in solution. When one of the proteins was mixed with POWT, the fluorescence emission spectrum increased in intensity and was blue shifted, indicating torsion of the POWT molecule in order to adjust to the protein block copolymers.

3.7.2 Co-assembly of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with different negatively charged polyelectrolytes

Both $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ co-assembled with the polyelectrolyte POWT, but POWT containing solutions and dispersions were not suitable for CD spectrometry. To see whether the mode of charge compensation (pH or polyelectrolyte) and whether the type of polyelectrolyte used influences the conformation of the protein block copolymer, we compared micrographs, CD measurements and SAXS measurements of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ at high pH with $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ at moderate pH but combined with either one of two different polyelectrolytes.

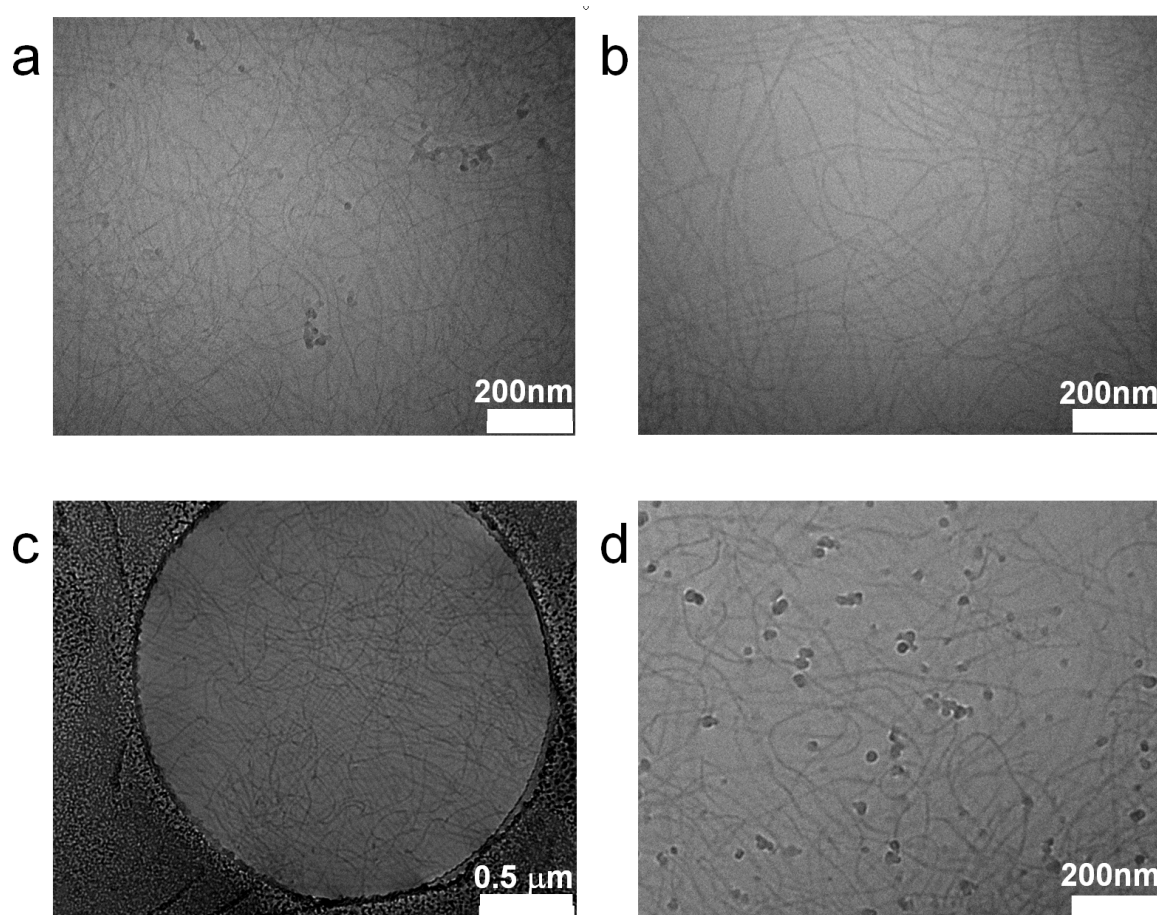


Figure 3.18 Cryo-TEM images of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ mixed with negatively charged polyelectrolyte to an f - of 0.5 and at pH 5.4 with (a, b) PAA (poly acrylic acid) and (c, d) $(\text{Zn-L}_2(\text{EO})_4)_n$. (Dark globules are ice crystals.)

The first negatively charged polyelectrolyte was the 42 monomer long poly acrylic acid (PAA) (**Fig. 3.1 b**). The second negatively charged polyelectrolyte was a low MW supramolecular polymer $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$, (**Fig 3.2**) composed of Zn^{2+} and a bisligand molecule consisting of two terdentate ligand (L^{2-}) groups (pyridine-2,6-dicarboxylic acid) connected at the 4-position of the pyridine ring by a spacer of four ethylene oxide (EO) units: $\text{L}_2(\text{EO})_4^{4-}$, (**Fig. 3.2 a**).

Divalent metal ions such as Zn^{2+} can be coordinated by two negative terdentate ligand groups. Overall such a complete coordination center carries two elementary negative charges, as illustrated in **Figure 3.2 b**. In aqueous solutions of a 1:1 mixture of Zn^{2+} and $\text{L}_2(\text{EO})_4^{4-}$, linear coordination polymers with a high degree of polymerization form spontaneously when the concentration is sufficiently high ($[\text{Zn}^{2+}] = [\text{L}_2(\text{EO})_4^{4-}] > 20 \text{ mM}$) [35, 36]. At low concentration, small coordination rings $\text{Zn}_2(\text{L}_2(\text{EO})_4)_2^{4-}$ are the dominant species (**Fig. 3.2 b**). These dilute rings can polymerize in the presence of positively charged polyelectrolyte, due to a cooperative process. We have reported an example of this [37].

Mixing the dissolved $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with other polyelectrolytes resulted in nanoscopic fibrils seen in the cryo-TEM micrographs both for $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with PAA (**Fig 3.18 a,b**) and for $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ (**Fig 3.18 c,d**). The cryo-TEM images of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ at high pH, with PAA and with $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ were taken on the same location with the same equipment. The fibrils at high pH gave very little contrast and were therefore imaged slightly out of focus to gain contrast. Therefore accurate dimensions for these fibrils could not be determined. The $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ fibrils with PAA and $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ gave much more contrast and could be imaged in focus. More contrast was due to more material per unit length of fibril, because co-assembly led to a self-assembled protein polymer core with an extra layer of polymer, opposed to pH induced fibrils that lack such an extra layer.

More contrast enabled us to obtain only the width of the $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ -PAA fibrils of about 12 nm, which could be consistent with a ribbon structure similar to the one described in **Figure 3.10**, but with an extra layer of polymer on each ribbon surface. From more detailed cryo-TEM micrographs (**Fig. 3.19**) taken on a different location, we obtained dimensions for the $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ - $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ fibrils. They appeared as long ribbons with a height of 2-4 nm and a width of 20-25 nm. The „edge on“ view of a ribbon (**Fig. 3.19 d** indicated with arrow **E**) is thin (2-4 nm) and has high contrast, because the viewed material is 20-25 nm thick (equal to the ribbon width). The „face on“ view (**Fig. 3.19 d** indicated with arrow **F**) is 20-25

nm broad and has less contrast because the ribbon is now viewed as a thin film with a thickness of only 2-4 nm. Mostly the ribbons are straight, but occasionally they can be observed to twist. This twist can be seen as a broad band with low contrast continuing as a thin band with high contrast, as indicated by the arrows in **Figure 3.19 c**.

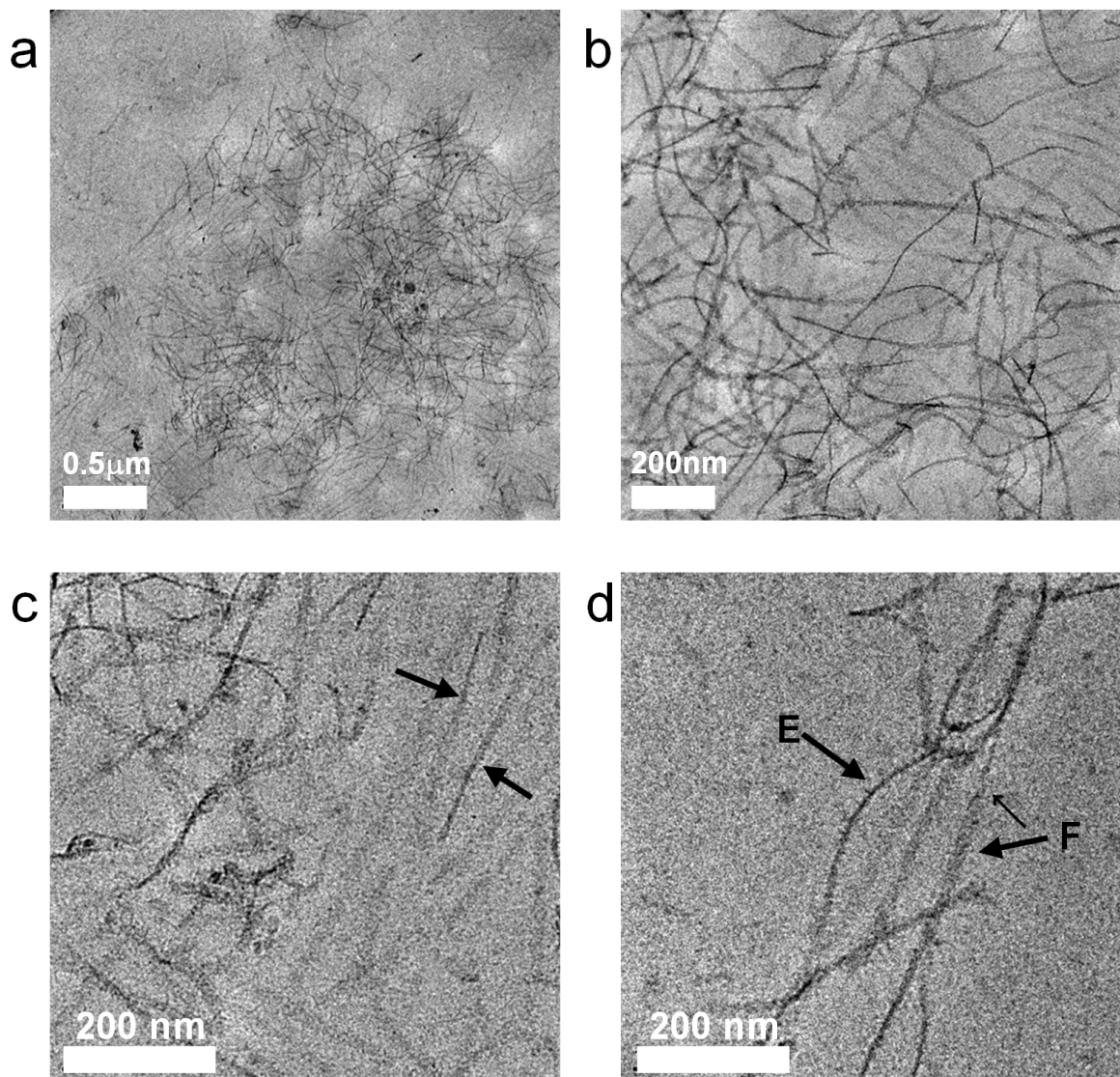


Figure 3.19 Cryo-TEM images of clearly visible ribbons from a mixture of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $(\text{Zn-L}_2(\text{EO})_4)_n$, at an f - of 0.5 at pH 5.4, in which the concentration of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ is 0.65 g/l. (a) Overview of ribbons. (b) Enlarged image of ribbons (c) Enlarged image of ribbons with twist indicated by arrows, where there is a transition from broad and low contrast ribbon to thin and high contrast ribbon. (d) Image of ribbons with thin, high contrast "edge on" view, indicated by arrow with E and broad, low contrast "face on" view indicated by arrow with F.

A new series of SAXS measurements was performed on $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ including the combination with PAA, and the combination with $(\text{Zn-L}_2(\text{EO})_4)_n$. For $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ in combination with PAA or $(\text{Zn-L}_2(\text{EO})_4)_n$, the measurements suggest anisotropic fiber-like

objects in which one fiber has cross-sectional dimensions of about 124 and 140 nm. These are probably stacks of ribbons that formed because the protein concentration used in the SAXS experiments was much higher than that used for cryo-TEM. The dimensions of individual ribbons could therefore not be determined with SAXS.

The occurrence of at least two different ribbons under similar conditions (all aqueous) may mean that there are different possible conformations of the S^H block leading to such differences in fibril dimensions. The width of the $(Zn-L_2(EO)_4^{2-})_n$ induced ribbons observed both in cryo-TEM and SAXS is consistent with a stack of β -sheets as reported earlier for crystals and ribbons from a methanol, formic acid mixture [26, 30], and not with a stack of β -rolls (**Fig. 3.8, 3.9**) as described above.

Because PAA and $(Zn-L_2(EO)_4^{2-})_n$ under various conditions had the same CD spectrum as the blank, these compounds did not influence the CD spectrum. Therefore, CD spectroscopy could be used to obtain indications of the structure of only the $S^H S^H$ blocks that self-assembled into fibrils, in response to the oppositely charged polyelectrolytes. After mixing, the observed spectra (**Fig. 3.20 a**) were different from those obtained at pH 11, in absence of PAA or $(Zn-L_2(EO)_4^{2-})_n$. The same subtraction of the CC spectrum from the $CS^H S^H C$ spectrum as described above was performed to obtain the spectrum of only the $S^H S^H$ block under different conditions (**Fig. 3.20 b**). The mix with PAA revealed a CD spectrum (**Fig 3.20 b**) that could not be deconvoluted well because of lack of appropriate reference spectra. Just as for **Figure 3.15 c,d**, either a new structure, different from the β -roll or a combination of structures, e.g. β -rolls and random structures is responsible for the obtained spectrum.

The mix with $(Zn-L_2(EO)_4^{2-})_n$ showed a CD spectrum (**Fig 3.20 b**) with positive ellipticity at 200 nm and negative minima at 210 nm and 220 nm. This spectrum is indicative of a β -sheet, although the molar ellipticity of the spectrum at 215 nm is smaller than one would expect for a β -sheet (at least $-5 \text{ kdeg} \cdot \text{dmol}^{-1} \cdot \text{cm}^2$ [38]), probably due to spectral flattening [32]. However, this spectrum is not that of a pure β -sheet and may be attributed to a combination of a combination of structures: β -sheet, α -helix, β -roll. In any case, CD spectrometry (**Fig 3.20 b**), in combination with cryo-TEM (**Fig 3.19**) indicates a well defined structure that is very different from the pure β -roll and from the combination of β -roll and random structures.

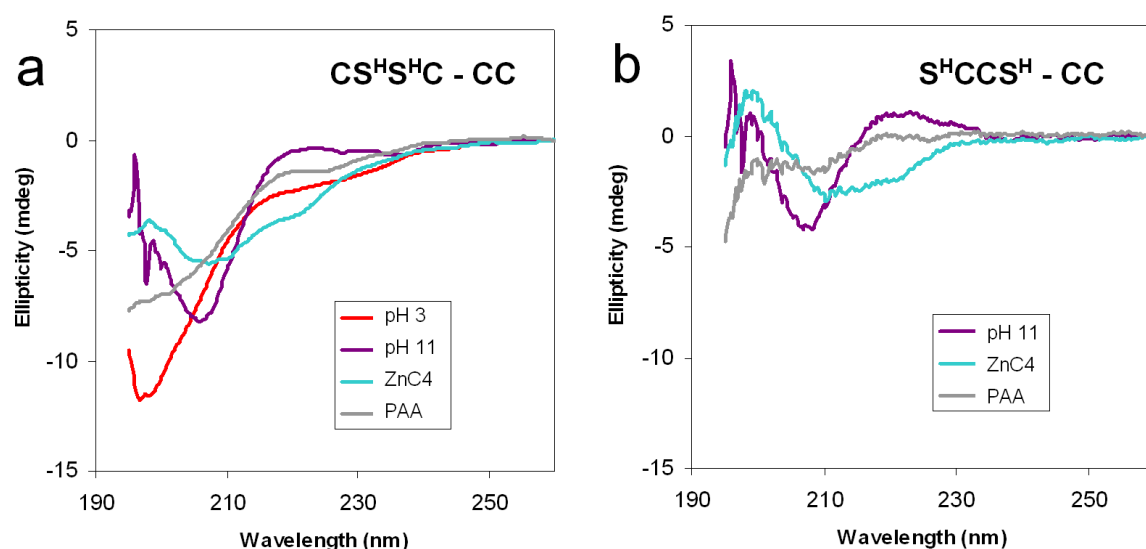


Figure 3.20 (a) The CD spectrum of charged $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ at pH 3 and the CD spectra of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ depending on the mode of charge compensation: by increasing pH, by adding $\text{Zn-L}_2(\text{EO})_4^{2-}$, and by adding PAA. (b) Analogous to figure 3.4, the spectrum of 0.042 g/l central $\text{S}^{\text{H}}\text{S}^{\text{H}}$ block by subtracting the 0.058 g/l CC spectrum from the 0.1 g/l $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ spectrum for the three different modes of charge compensation.

Let us summarize the structures that we observed for $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ under different conditions. At low pH, $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ forms a random and extended structure. At high pH it forms a mix of β -roll ribbons, and random, kinetically trapped structures. $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ co-assembles with PAA to form either a combination of random structures and well defined ribbons with a core width of about 12 nm, consistent with a core of stacked β -rolls, or, another orderly but still unresolved structure with the same core width. The $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ induced $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ co-assembly contains a $\text{S}^{\text{H}}\text{S}^{\text{H}}$ ribbon core with a very different structure: either a combination of β -sheet α -helix and β -roll or another unresolved structure.

The occurrence of such conformational differences under such similar conditions, indicates that all of these structures are energetically very close to each other, and might be interconvertable.

3.7.3 Co-assembly of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ or $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ with $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ or $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$

Ultimately, the histidine and the glutamic-acid containing protein block copolymers were created to have a set, complimentary in charge, to co-assemble at moderate pH where both molecules are charged. This would form an all protein co-assembly suitable for application at physiological conditions. When 8 g/l protein solutions were mixed, in a ratio of

1:1, all combinations of positively with negatively charged protein block copolymer produced a gel. To investigate the structure of these gels we used CD spectroscopy and TEM.

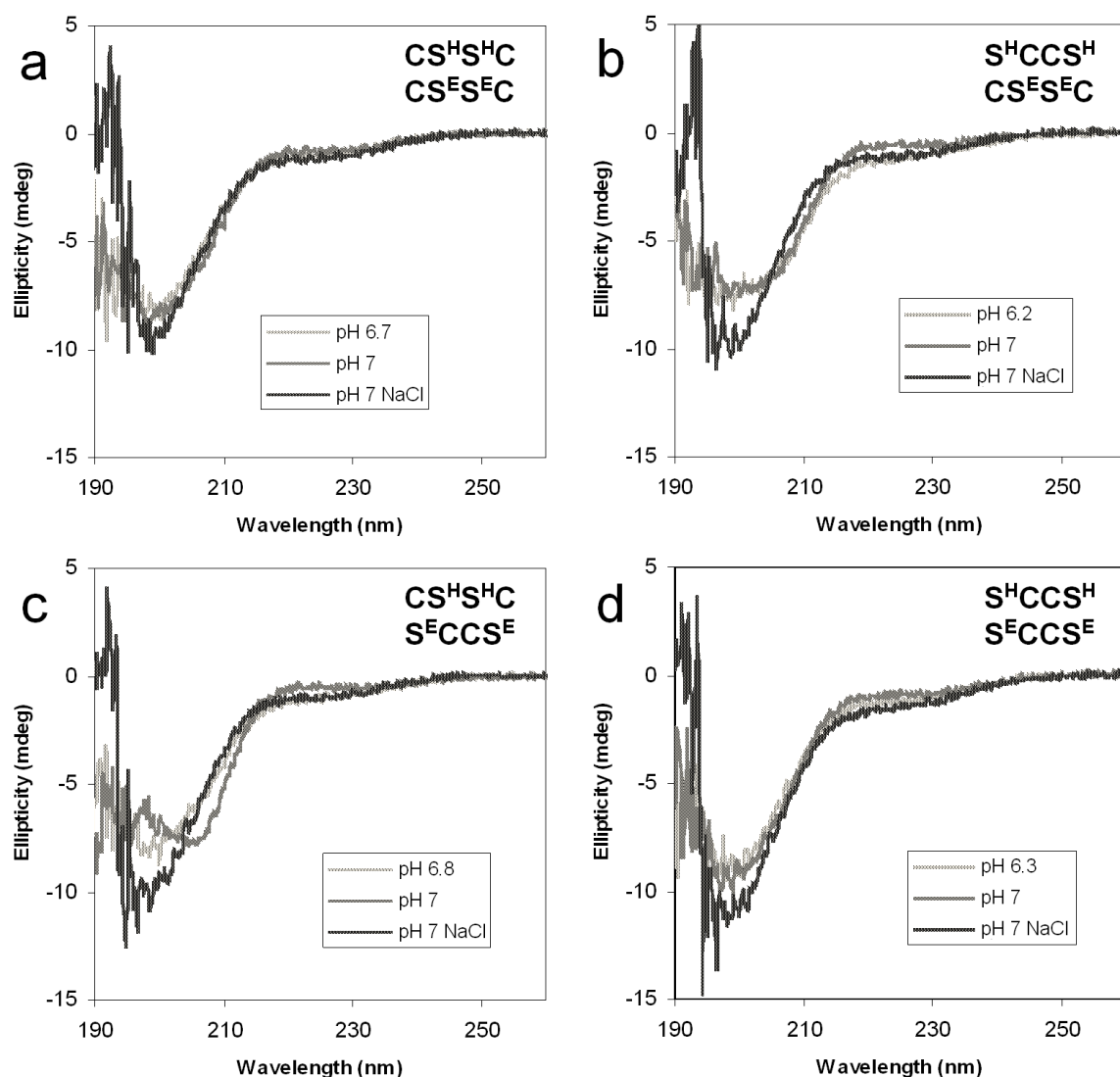


Figure 3.21 For all possible 0.1 g/l, 1:1 mixtures of oppositely charged protein block copolymers, the following three CD spectra: the unbuffered mix, the mix buffered with 5 mM phosphate at pH 7 and the mix at pH 7 with 50 mM NaCl. The depicted spectra are from at least 17 h and at the most, 23h after mixing. pH varied from 6.2 to 6.8 in the unbuffered mixes.

CD spectra were taken for all 4 possible combinations of a negatively charged protein polymer with a positively charged protein polymer (**Fig. 3.21**). This was done at different times for three different conditions: 1) unbuffered, 2) 5 mM phosphate buffer, and 3) 5 mM phosphate buffer containing 50 mM NaCl. For all combinations, the CD spectrum of the unbuffered and buffered mix did not change after 18 h. Just after mixing, the CD spectra of

both the buffered and unbuffered mixes were equal to that of the salt containing mix at larger times. For every mix, the three CD spectra are shown in **Figure 3.21**.

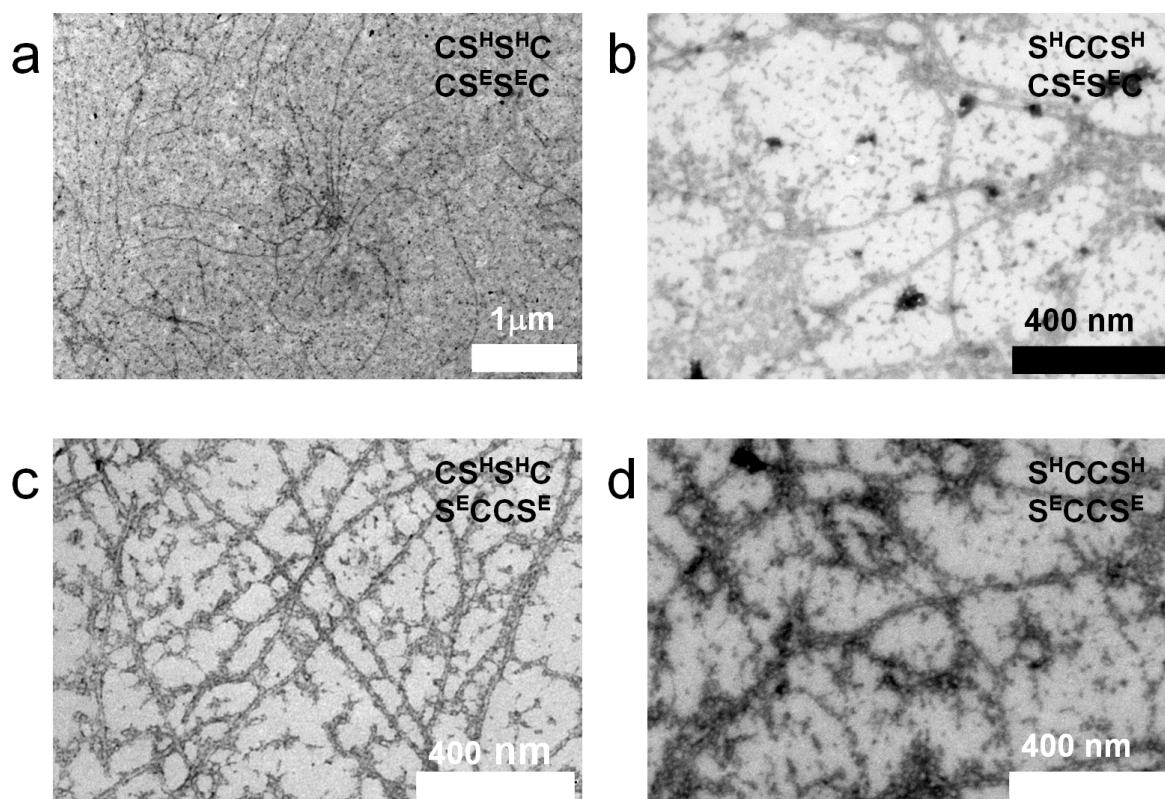


Figure 3.22 TEM images of all possible unbuffered 0.1 g/l 1:1 mixtures of the produced oppositely charged protein block copolymers.

The CD spectra indicate hardly any conformational change upon mixing of the separate, oppositely charged, protein block copolymers. All polymers tend to persist in their random structure. All spectra shown in **Figure 3.21** are from samples that were allowed to age after mixing for more than 17 h. The spectra for the mixes with 50 mM NaCl are practically the same as for the mixes immediately after mixing (not shown). For the mix of CS^HS^HC and CS^ES^EC (**Fig. 3.21 a**) there was no change in CD spectrum at all. For the mix of S^HCCS^H and S^ECCS^E (**Fig. 3.21 d**) the change in CD spectrum was minimal and might just as well have been caused by some spectral flattening [32]. For the mix of S^HCCS^H and CS^ES^EC (**Fig. 3.21 b**) after 18 h, we see that both the buffered and unbuffered mix show a decrease and a broadening of the minimum at 200 nm, shifting to slightly larger wavelengths. This does indicate a minor conformational change. Most change, is observed for the mix of CS^HS^HC and S^ECCS^E (**Fig. 3.21 c**), where, after 18 h, a maximum at 198 nm and a minimum at 204 nm can be observed for the buffered mix. This maximum and minimum is reminiscent of the similar

spectra that all of our protein polymers display as a single component, when forming fibrils at an extreme pH, and could therefore be caused by a conformational change in either component.

Since we do observe a gel when mixing these polymers at higher concentrations, but hardly see conformational change when measuring CD spectra of mixes, we must conclude that aggregation does occur but that structures stay random and may be kinetically trapped. Adding 50 mM NaCl in the solution seems to further suppress the formation of defined structures leading to even less CD spectral change when mixing the different oppositely charged products.

The same solutions used for CD spectroscopy were used for TEM imaging. In general, the observations for all combinations could be described as random aggregates with some linear aggregates. For the combination of CS^HS^HC with CS^ES^EC (**Fig. 3.22 a**) some linear aggregates can be observed in what appears to be a “sea” of monomers or maybe some random network. For the other mixes (**Fig. 3.22 b-d**) we see what appears to be lots of pearls and some strings of pearls. The micrograph with the largest amount of strings of pearls, compared to the amount of randomly organized pearls is **Figure 3.22 c**, which coincides with **Figure 3.21 c** in which the largest CD effect of mixing was observed. The strings of pearls could be interpreted as the oppositely charged polymers stacking successively into a linear aggregate, however they might also be interpreted as a fibril core of the one polymer coated by a random structure of the other. In retrospect, both the drying and the pH of the staining agent (uranyl acetate) could have influenced the structures that we see in TEM. For a more insight into the structures that we studied with CD spectrometry, cryo-TEM should have been performed.

3.8 Conclusion Section III

Both CS^ES^EC and CS^HS^HC co-assembled with oppositely charged polyelectrolytes, for example the zwitterionic POWT, to form fibrils. This suggests the use of these protein block copolymers to template fibrillar structures onto other (oppositely charged) polymers that would by themselves not form fibrils, but have properties that are desired in a defined fibrillar form. For example a conductive polymer like POWT, which might be combined with CS^HS^HC to form conductive composite ribbons for nanowires. However, both in combination with CS^ES^EC and CS^HS^HC, a blue shift of the emission spectrum was observed for POWT,

indicating torsion of the molecule which might lead to decreased conductivity of this polymer. Still, other conductive polyelectrolytes may be used to this end.

Charge compensating $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with different negatively charged polyelectrolytes or pH, led to various self- or co-assembled ribbon core forms, associated with different conformations of the $\text{S}^{\text{H}}\text{S}^{\text{H}}$ block. High pH led to a mixture of ribbons and random structures, with a CD spectrum indicating a structure similar to that of $\text{S}^{\text{E}}\text{S}^{\text{E}}$ at low pH, possibly a β -roll, and to that reported previously for YEHK [29]. Mixing $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ with PAA led to probably two different structures in one sample reflected by cryo-TEM images and CD spectroscopy: ribbons with a width of about 12 nm and probably random structures. Adding low MW supramolecular polymer $(\text{Zn-L}_2(\text{EO})_4^{2-})_n$ led to efficient self-assembly of $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ into ribbons that may contain a combination of β -sheet with α -helix and β -roll. It seems that different conformations are possible under similar conditions, suggesting that these different conformations are energetically close to each other and may be interconvertable by shifting physical conditions.

Mixing the different oppositely charged protein block copolymers led to gels at higher protein concentrations but, according to CD spectrometry, hardly to any conformational changes. In TEM images, which is not completely representative of the situation in aqueous environment, we observed both randomly and linearly aggregated pearl-like objects. Possibly a fraction of one component self-assembles into ribbons while the other covers the surface of the self-assembly to compensate the charge.

3.9 Summary and conclusion of chapter 3

At neutral pH all four products $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$, $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$, $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ are water soluble in their pure form. All four products form fibrils when the charge of the silk-like S blocks is compensated by changing the pH. Both $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ form fibrils when their charge is compensated by oppositely charged polyelectrolyte. Also $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ are likely to do so. Fibril formation is aberrant for triblock copolymers which usually form micelles, or flower-like micelles, and molecular networks, depending on whether the middle block or the outer blocks aggregate. The formation of fibrils by these protein polymers means that the aggregating blocks must be highly structured and organized.

	$CS^E S^E C$	$S^E C C S^E$
Low pH	Ribbons β -rol core 11.6 x 2.7 nm Hydrophilic corona Single filaments (dry)	Ribbons β -rol core 12.1 x 2.7 nm Hydrophilic corona loops Bundles of filaments (dry)
POWT	Filaments (nano-sized)	ND
	$CS^H S^H C$	$S^H C C S^H$
pH high	Combination of nano- ribbons and globules	Combination of nano- ribbons and globules
POWT	Filaments (nanosized)	ND
PAA	Ribbons 12 nm wide core structure unknown	ND
Zn C4	Ribbons core 23 x 3 nm structure unresolved	ND
	$CS^H S^H C$	$S^H C C S^H$
$CS^E S^E C$	Mainly random and some linear aggregates	Random aggregates with “strings of pearls “
$S^E C C S^E$	Random aggregates with “strings of pearls “	Random aggregates with “strings of pearls “

Table 3.1 Overview of different self- and co-assembled structures formed by the four studied different protein block copolymers under, different conditions, and in combination with different oppositely charged polyelectrolytes including each other.

A core shell structure is most likely for the S^E containing self-assemblies at low pH, just as it is probable for all the other ribbon containing self- and co-assemblies described in this chapter. However, a core shell structure still has to be proven experimentally for all of them. Assuming such a structure, the core is composed of the silk-like S block which folds and stacks compactly into a well defined fibril with internal hydrogen bonds and apolar domains, filled with the methyl side groups of the alanine residues of the silk-like S blocks. The shell is made of the hydrophilic water swollen random coil Collagen-like C block of which form nor

structure varies with conditions. All SCCS fibrils probably have corona of CC block loops because both S end blocks of the same molecule sit in the same fibril core, while CSSC has a corona of loose C block ends. In contrast to the highly structured, self-assembled fibril core, the morphology of the C block is a feature that our products have in common with classical block copolymers. When they have aggregating end blocks, they form flowerlike micelles with loops on the surface comparable the loops on our SCCS ribbons and if they have aggregating middle blocks, they form micelles with a corona of ending chains.

Several factors influence the conformation of the uncharged S block and therefore also the form of the fibril core. We found that both core form and S block conformation depend on the sample history (solvent), on mechanisms for charge compensation, and type of polyelectrolyte used, and also on the amino acid residue in the X position of the (GAGAGAGX)_n repeat. The different structures were all aqueous and triggered to form with relatively mild stimuli. Therefore the different (aqueous) structures have a similar free energy and may be interconvertable.

Let us summarize the structures found. Apart from the charged random structure, that the S blocks can assume, we found 3 different structures for the fibril core (**Table 3.1**). The first structure found was produced by both CS^ES^EC and S^ECCS^E, when acidified. They produced micron long ribbons with cross-sectional dimensions of 2.7nm and 12 nm, possibly made from a stack of unusual β -rolls that contain the glutamic acids in the turns. For CS^HS^HC and S^HCCS^H at high pH, a population of molecules probably assumes the same β -roll and ribbon structure, while an other part of the molecules forms a network of kinetically trapped random structures. The CS^HS^HC and S^HCCS^H samples that contain ribbons, have CD spectra similar to those of the S^E-ribbon containing polymers at low pH. As explained above for all four protein block copolymers, SCCS order resulted in CC loops and larger core dimensions than CSSC order.

Similarly to at high pH, CS^HS^HC of which charge was compensated with PAA, self-assembled into ribbons with a width of about 12 nm. The CD spectrum differed from that at high pH in such a way that the structure causing it may either be a combination of β -roll and other e.g. random structures or a different structure altogether.

The second structure found could be a stack of β -sheets, similar to that reported earlier for crystals and ribbons from a methanol, formic acid mixture [26, 30] but then in combination with a small amount of β -roll. CD spectroscopy pointed to this combination and

the width of the ribbons is consistent with a stack of β -sheet forming S blocks, with the histidine residues in the turn.

The third and last structure is the pearls, and strings of pearls that we found when combining either $CS^E S^E C$ or $S^E CCS^E$ with either $CS^H S^H C$ or $S^H CCS^H$. In CD spectral measurements, the conformation can be seen to stay mainly random while a fraction of the proteins may form fibrils, either by successive stacking of positively and negatively charged proteins, or by self-assembly of only one component while the other component coats the fibril core of the first component, neutralizing its charge.

For the POWT co-assemblies, only fibril formation has been demonstrated, but whether it is one of the structures mentioned above, or whether it is yet another structure is unknown.

Both $CS^E S^E C$ and $CS^H S^H C$ co-assembled to form ribbons with oppositely charged polyelectrolytes at moderate pH. This suggests the use of these protein block copolymers to template nano-ribbons of other (oppositely charged) polymers that would by themselves not form nano-ribbons, but have desirable properties for them. For example conductive polymers, like POWT, might be complexed into a composite conductive nanowire. The formation of co-assembled fibrils and gels at moderate pH also opens up possibilities for applications at physiological conditions like scaffolds for tissue engineering.

Acknowledgements

The authors thank Luben Arnaudov, Remco Fokink and Adrie Westphal for sharing their knowledge and for instructions on atomic force microscopy (AFM) imaging, dynamic light scattering (DLS), and CD spectrometry respectively. The authors also thank Adriaan van Aelst for the TEM images and Marcus Drechsler for the cryo-TEM images

References

- 1 Woolfson, D. N. and Ryadnov, M. G. (2006) Peptide-based fibrous biomaterials: some things old, new and borrowed. *Current Opinion In Chemical Biology* **10**, 559-567
- 2 Chen, C. S., Mrksich, M., Huang, S., Whitesides, G. M. and Ingber, D. E. (1997) Geometric control of cell life and death. *Science* **276**, 1425-1428
- 3 Chitkara, D., Shikanov, A., Kumar, N. and Domb, A. J. (2006) Biodegradable injectable in situ depot-forming drug delivery systems. *Macromolecular Bioscience* **6**, 977-990
- 4 Harrington, D. A., Cheng, E. Y., Guler, M. O., Lee, L. K., Donovan, J. L., Claussen, R. C. and Stupp, S. I. (2006) Branched peptide-amphiphiles as self-assembling coatings for tissue engineering scaffolds. *Journal Of Biomedical Materials Research Part A* **78A**, 157-167
- 5 Krausch, G. and Magerle, R. (2002) Nanostructured thin films via self-assembly of block copolymers. *Advanced Materials* **14**, 1579-1583
- 6 Muthukumar, M., Ober, C. K. and Thomas, E. L. (1997) Competing interactions and levels of ordering in self-organizing polymeric materials. *Science* **277**, 1225-1232

- 7 Park, M., Harrison, C., Chaikin, P. M., Register, R. A. and Adamson, D. H. (1997) Block copolymer lithography: Periodic arrays of similar to 10(11) holes in 1 square centimeter. *Science* **276**, 1401-1404
- 8 Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A. and Stupp, S. I. (2004) Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science* **303**, 1352-1355
- 9 Mori, H. and Muller, A. H. E. (2003) New polymeric architectures with (meth)acrylic acid segments. *Progress in Polymer Science* **28**, 1403-1439
- 10 Winnik, M. A. and Yekta, A. (1997) Associative polymers in aqueous solution. *Current Opinion in Colloid & Interface Science* **2**, 424-436
- 11 Werten, M. W. T., Wisselink, W. H., van den Bosch, T. J. J., de Bruin, E. C. and de Wolf, F. A. (2001) Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Engineering* **14**, 447-454
- 12 Werten, M. W. T., Moers, A. P. H. A., Vong, T., Zuilhof, H., van Hest, J. C. M. and de Wolf, F. A. (2008) Biosynthesis of an Amphiphilic Silk-Like Polymer. *Biomacromolecules*
- 13 Andersson, M., Ekeblad, P. O., Hjertberg, T., Wennerstrom, O. and Inganas, O. (1991) Polythiophene with a Free Amino-Acid Side-Chain. *Polymer Communications* **32**, 546-548
- 14 Vermonden, T., van der Gucht, J., de Waard, P., Marcelis, A. T. M., Besseling, N. A. M., Sudholter, E. J. R., Fleer, G. J. and Stuart, M. A. C. (2003) Water-soluble reversible coordination polymers: Chains and rings. *Macromolecules* **36**, 7035-7044
- 15 Yun Yan, Aernout A. M. Nicolaas A. M. B. Frits A. de W. A. de K. M. D. Martien A. Cohen S. (2008) Nanoribbons Self-Assembled from Triblock Peptide Polymers and Coordination Polymers13. *Angewandte Chemie International Edition* **47**, 4192-4195
- 16 Van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. and Berendsen, H. J. C. (2005) GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **26**, 1701-1718
- 17 Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *Journal of Physical Chemistry B* **105**, 6474-6487
- 18 Hess, B., Bekker, H., Berendsen, H. J. C. and Fraaije, J. (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472
- 19 York, D. M., Darden, T. A. and Pedersen, L. G. (1993) The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. *Journal of Chemical Physics* **99**, 8345-8348
- 20 Frenkel, D. and Smit, B. (2002) *Understanding Molecular Simulation*
- 21 Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *Journal of Molecular Graphics* **14**, 51-&
- 22 Fossey, S. A., Nemethy, G., Gibson, K. D. and Scheraga, H. A. (1991) Conformational Energy Studies of Beta-Sheets of Model Silk Fibroin Peptides .1. Sheets of Poly(Ala-Gly) Chains. *Biopolymers* **31**, 1529-1541
- 23 Asakura, T., Suita, K., Kameda, T., Afonin, S. and Ulrich, A. S. (2004) Structural role of tyrosine in Bombyx mori silk fibroin, studied by solid-state NMR and molecular mechanics on a model peptide prepared as silk I and II. *Magnetic Resonance in Chemistry* **42**, 258-266
- 24 Krieger, E. (2003) YASARA.
- 25 Borsboom, M., Bras, W., Cerjak, I., Detollenaere, D., van Loon, D. G., Goedtkindt, P., Konijnenburg, M., Lassing, P., Levine, Y. K., Munneke, B., Oversluizen, M., van Tol, R. and Vlieg, E. (1998) The Dutch-Belgian beamline at the ESRF. *Journal of Synchrotron Radiation* **5**, 518-520
- 26 Smeenk, J. M., Otten, M. B. J., Thies, J., Tirrell, D. A., Stunnenberg, H. G. and van Hest, J. C. M. (2005) Controlled assembly of macromolecular beta-sheet fibrils. *Angewandte Chemie-International Edition* **44**, 1968-1971
- 27 Topilina, N. I., Higashiya, S., Rana, N., Ermolenkov, V. V., Kossow, C., Carlsen, A., Ngo, S. C., Wells, C. C., Eisenbraun, E. T., Dunn, K. A., Lednev, I. K., Geer, R. E., Kaloyeros, A. E.

- and Welch, J. T. (2006) Bilayer fibril formation by genetically engineered polypeptides: Preparation and characterization. *Biomacromolecules* **7**, 1104-1111
- 28 Crisma, M., Fasman, G. D., Balaram, H. and Balaram, P. (1984) Peptide Models for Beta-Turns - a Circular-Dichroism Study. *International Journal of Peptide and Protein Research* **23**, 411-419
- 29 Lednev, I. K., Ermolenkov, V. V., Higashiya, S., Popova, L. A., Topilina, N. I. and Welch, J. T. (2006) Reversible thermal denaturation of a 60-kDa genetically engineered beta-sheet polypeptide. *Biophysical Journal* **91**, 3805-3818
- 30 Krejchi, M. T., Cooper, S. J., Deguchi, Y., Atkins, E. D. T., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Crystal structures of chain-folded antiparallel beta-sheet assemblies from sequence-designed periodic polypeptides. *Macromolecules* **30**, 5012-5024
- 31 Cantor, E. J., Atkins, E. D. T., Cooper, S. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Effects of amino acid side-chain volume on chain packing in genetically engineered periodic polypeptides. *Journal Of Biochemistry* **122**, 217-225
- 32 Bustamante, C. and Maestre, M. F. (1988) Statistical Effects in the Absorption and Optical-Activity of Particulate Suspensions. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 8482-8486
- 33 Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) 3-Dimensional Structure of the Alkaline Protease of *Pseudomonas-Aeruginosa* - a 2-Domain Protein with a Calcium-Binding Parallel-Beta Roll Motif. *Embo Journal* **12**, 3357-3364
- 34 Chen, C. C., Krejchi, M. T., Tirrell, D. A. and Hsu, S. L. (1995) Effect of Water on the Structure of a Model Polypeptide. *Macromolecules* **28**, 1464-1469
- 35 Vermonden, T., Branowska, D., Marcelis, A. T. M. and Sudholter, E. J. R. (2003) Synthesis of 4-functionalized terdentate pyridine-based ligands. *Tetrahedron* **59**, 5039-5045
- 36 Vermonden, T., van Steenbergen, M. J., Besseling, N. A. M., Marcelis, A. T. M., Hennink, W. E., Sudholter, E. J. R. and Stuart, M. A. C. (2004) Linear rheology of water-soluble reversible neodymium(III) coordination polymers. *Journal of the American Chemical Society* **126**, 15802-15808
- 37 Yan, Y., Besseling, N. A. M., de Keizer, A., Drechsler, M., Fokkink, R. and Stuart, M. A. C. (2007) Wormlike aggregates from a supramolecular coordination polymer and a diblock copolymer. *Journal of Physical Chemistry B* **111**, 11662-11669
- 38 Ganesh, S. and Jayakumar, R. (2003) Structural transitions involved in a novel amyloid-like beta-sheet assemblage of tripeptide derivatives. *Biopolymers* **70**, 336-345

4

Nucleation and growth by a silk-collagen-like block copolymer into supramolecular nano-ribbons

A.A. Martens, G. Eggink, F.A. de Wolf, M.A. Cohen Stuart

Abstract

From earlier chapters, we may conclude that the macroscopic gels formed by the protein block copolymer $CS^E S^E C$ consist of nano-ribbons, and that ribbon formation goes hand in hand with conformational change that can be followed with CD spectroscopy. CD spectroscopy at 200 nm was used in this study to follow the ribbon formation of $CS^E S^E C$ in time, in dilute, acidified solutions, revealing a nucleation and growth mechanism for $CS^E S^E C$ ribbons under a critical pH of approximately 4.5. When the pH was increased, the ribbons dissolved and the formed gels melted, but they only did so above pH 5.4 which is much higher than the critical pH of ribbon formation. This pH region in which ribbons do not form, nor dissolve suggests a kinetic barrier to ribbon formation. The purified and freeze-dried $CS^E S^E C$ appeared to contain nuclei from which ribbon growth could start.

4.1 Introduction

In **Chapter 2**, we describe the production and purification of silk-collagen-like block copolymers S^ECCS^E and CS^ES^EC which can form gels when the negative charge of the S^E block is neutralized by lowering the pH. In this chapter we consider the rate of the various processes leading to gelling by collecting time dependent data after a drop in pH.

The formation rate of the ribbons that constitute the gels is of great interest with respect to applications and material processing. During aggregation and before setting of the gel, the dispersion still flows and can be cast in any form. Different applications have different kinetic requirements, For example, if living cells were to be cast into such a gel, the gel has to solidify within a certain time, because of biological requirements.

Future processing of the gel into materials via techniques like casting or gel spinning have to occur while the molecules are busy aggregating but the gel is not matured and still pliable. Insite in the mechanisms behind the aggregation behavior of our protein block copolymers might help tuning aggregation behavior for better material processing.

Self-assembly of CS^ES^EC molecules in solution is triggered by lowering the pH. In doing so, the S^E blocks self-assemble into nano-sized ribbons, as was described in **Chapter 3**. The conformational changes in response to changing pH are reflected in the CD spectrum. As was discussed in **Chapter 2**, lowering the pH, first caused an almost instant change in CD spectrum, after which a second, slow CD spectral change occurred over several hours while ribbons started to appear. We proposed that the quick change in CD spectrum was due to quick charge neutralization, loss of self-repulsion and, random rearrangement of the S^E block. We attributed the slow CD spectral change, to CS^ES^EC molecules folding into the conformation present in ribbons. The largest change in ellipticity was observed at 200 nm. Therefore, we will use the ellipticity at 200 nm to monitor conformational change and ribbon formation for many samples.

Ribbon formation of proteins and synthetic polypeptides has been the subject of kinetic studies before. The proteins involved in these studies were natural or trangenically produced proteins for medical research [1] or industrial research [2] purposes. Protein fibril growth generally displayed nucleation and growth kinetics, sometimes with multiple intermediates between monomer and fibril. Small synthetic peptides, designed to self-assemble based on β -sheet formation [3-6], generally nucleate and self-assemble quicker than the larger proteins,

when induced to form fibrils. Structure formation kinetics of molecules comparable to ours[7-10] have not been studied to date.

4.2 Materials and methods

4.2.1 Protein block copolymers

The synthesis of the protein block copolymers $CS^E S^E C$, $S^E CCS^E$, and their primary amino acid sequences have been described previously in **Chapter 2**. Unless mentioned otherwise, stock solutions contained 1 g/l of product in 1 mM NaOH or HCl. $CS^E S^E C$ and $S^E CCS^E$ were prepared in NaOH and $CS^H S^H C$ and $S^H CCS^H$ were prepared in HCl.

4.2.2 CD spectrometry

Samples of 0.1 and 0.2 g/l $CS^E S^E C$ were prepared at different pH by mixing a 1 g/l $CS^E S^E C$ stock solution with MQ water and HCl in a total volume of 2 ml. 400 μ l was transferred to a 1 mm path length quartz cuvette and pH was measured in the remaining 1.6 ml. For kinetic studies, ellipticity at 200 nm was recorded every 20 seconds for several hours at 21°C (**Fig. 4.1**) using a Jasco J-715 spectropolarimeter equipped with a Peltier temperature controller. A monoexponential function $Y=B-A*e^{(-C*t)}$ was fitted every CD data series (**Fig. 4.1**) according to the least squares method, using the MS Excel solver function. from the monoexponential function, amplitude (A) and initial slope ($A*C$) were derived and used for further calculations.

4.2.3 Gel swelling

A gel drop was formed in the severed cap of an Eppendorf tube by mixing 200 μ l of 20 g/l $CS^E S^E C$ dissolved in 10 mM NaOH with 6 μ l of 1M HCl. After 20 minutes the gel had formed and the combined weight of the cap and the gel was determined. The weight of the cap had been determined before the experiment and the rest of the tube was used as an airtight cover to prevent evaporation and the whole was kept upside down. The gel drop was allowed to equilibrate with 200 μ l of water for one week, after which most water was removed with a micropipette, and the remaining water was removed carefully with a paper tissue. The combined weight of the gel drop and cap were determined again. The drop was left to dry for 14 hours after which the combined gel drop and cap were weighed once more. The gel drop was then resubmerged in the removed solution for 14.5 hours, after which the solution was

removed, as described above, and the combined gel drop and cap were weighed. The solution was added again and after another 9.5 hours the solution was removed again and the final combined weight of the drop and cap was determined.

4.2.4 pH melting point determination

To make a stock solution, 1 g/l freeze-dried CS^ES^EC was dissolved in 1 mM NaOH. To every of 10 cuvettes was added: 900 μ l CS^ES^EC stock solution and 100 μ l 100 mM HCl. The mixture was allowed to gel for 1 h at 70°C. On top of the gels, 100 mM phosphate buffer was added; buffer pH was chosen such that the final pH in the samples ranged from 5.2 to 5.7. Gel and buffer were kept at 70°C for another hour to eliminate bacterial contamination. The buffers were allowed to penetrate into the gel for six weeks at room temperature. Subsequently DLS was measured on a home-built fixed angle light scattering setup using an ALV5000 digital correlator. A neutral density filter (OD=0.8) was used to adjust the intensity of the 532 nm DPSS laser as to match the linear range of the photo multiplier tube. For each cuvette the intensity autocorrelation $g_2(\tau)$ function was measured, from which $\Delta g_2 = g_2(0) - g_2(\infty)$ was calculated (ref art 1), and plotted against pH (**Fig. 4.6**). After DLS measurements, pH was measured. For each sample it was found to be the same in the liquid as in the gel.

4.3 Results

4.3.1 CD spectra reflecting conformation change

In **Chapter 3 (Fig. 3.5 a)**, circular dichroism (CD) spectra of a 0.1 g/l CS^ES^EC solution were determined, reflecting its conformational response to lowering the pH. The response in CD signal had the largest amplitude at 200 nm. Therefore we followed the conformational change in time for this protein polymer in different solutions, by following the ellipticity at 200 nm.

The conformational change leading to a change in ellipticity was described in **Chapter 2** to occur simultaneously with ribbon formation, and so we assumed that we could monitor the ribbon formation by following the CD signal at 200 nm. **Figure 4.1 a** is an example of such a measurement for 0.1 g/l CS^ES^EC at pH 2.48 and 21°C, with time after acidification (h) on the abscissa and ellipticity (mdeg) on the ordinate. In the same way, ellipticity at 200 nm was followed as a function of time for several samples.

Using the least squares method, CD data for every measured sample was fitted with a monoexponential curve $Y=B-A \cdot e^{(-C \cdot t)}$ from which the parameters amplitude (A) and initial

slope ($A \cdot C$) were extracted. **Figure 4.1 a** shows the monoexponential curve fitted through the data points, suggesting first order kinetics for ribbon formation.

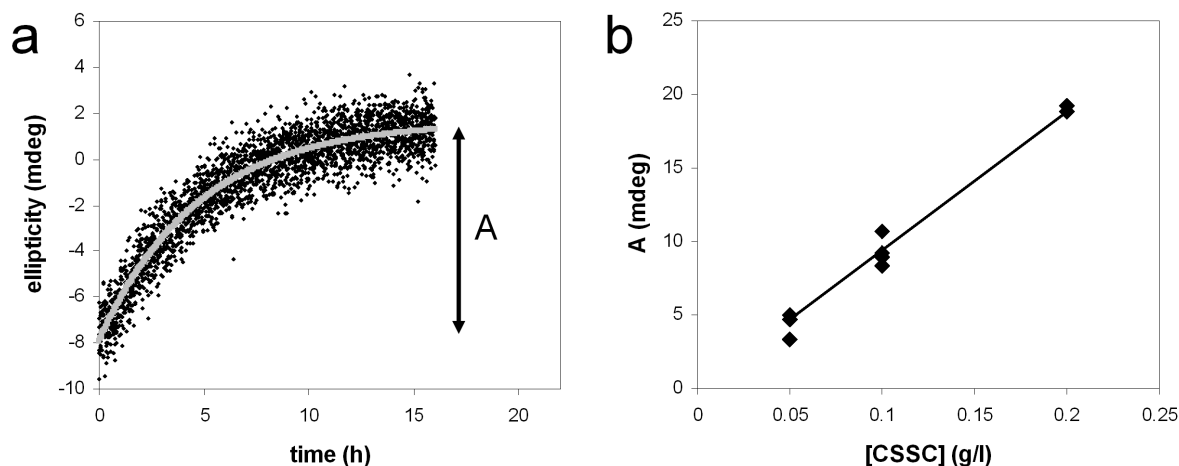


Figure 4.1 (a) CD data points at 200 nm for a 0.1 g/l $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ solution at pH 2.48 and 21°C, the gray line being the corresponding fitted monoexponential curve as a function of time after acidification, with amplitude A indicated by a double headed arrow. (b) Amplitude of the CD signal change, derived from monoexponential curves fitted to time resolved CD measurements as a function of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentration. This amplitude is directly proportional to the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentration.

4.3.2 Quantitative relationship between CD_{200} and ribbon concentration

For samples at pH 1.5 both at 15°C and at 21°C, the amplitude of the ellipticity at 200 nm (A) was plotted as a function of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentration (**Fig. 4.1 b**). The line through the data points, passes through the origin, indicating the absence of a lower critical $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentration, under which there is no conformational change. Since, according to **Chapter 2**, the conformational change occurring in the course of about 10-30 h goes hand in hand with ribbon formation, we assume that when the exponential rise of the CD signal nears completion, so does the ribbon formation.

The measured amplitude of the CD signal at 200 nm is directly proportional to the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentration (**Fig. 4.1 b**). Assuming that all of the material in the sample is completely converted into ribbons at the end of the process, **Figure 4.1 b** allows to calibrate the CD amplitude against the ribbon concentration, the conversion factor being 94 mdeg $\cdot\text{gl}^{-1}$. The further assumption that the amplitude reached at any given time with respect to the CD signal at $t=0$ is proportional to the fibril concentration at this given time, enables the calculation of the of the fibril concentration at any given time from curves fitted to CD data.

4.3.3 Initial folding rate as a function of pH and CS^ESC concentration

For several samples varying in pH and CS^ESC concentration, initial slopes of the time resolved CD measurements at 200 nm (mdeg*h⁻¹) were converted to initial folding rates of the protein into ribbon (g l⁻¹h⁻¹) by using the conversion factor calculated above. These initial folding rates plotted as a function of pH are given in **Figure 4.2 a**. Above pH 4.5 = pH*, no ribbon formation takes place. The reaction rate increases with decreasing pH, starting from pH ≈ 4.5. Surprisingly, two different dependencies of initial folding rate on pH were found for two separate CS^ESC stock solutions, both prepared in the same fashion. Both stock solutions contained 10 g l⁻¹ CS^ESC and 10 mM NaOH. The only difference was that stock 1, had been freshly prepared from freeze-dried material and tests were done without freezing the stock between measurements, whereas stock 2, which had also been prepared from the same freeze-dried material, was older and had been frozen and thawed several times before it was used to prepare samples for CD measurements. Samples prepared from the fresh stock 1 seem to form ribbons more rapidly than samples prepared from the older stock 2.

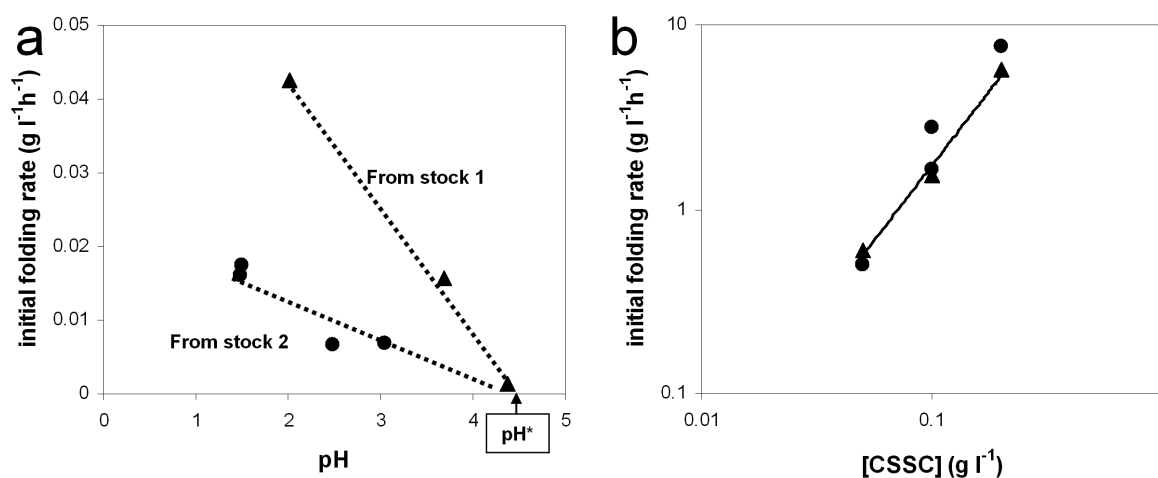


Figure 4.2 (a) Initial slopes of the ribbon forming reaction, according to time resolved CD measurements, as a function of pH, for a 0.1 g/l CS^ESC solution prepared from 1: a freshly prepared 1 g/l stock solution in 1 mM NaOH and 2: an other stock solution prepared in the same way but aged. **(b)** In samples prepared from stock 2: initial slope of the ribbon forming reaction, according to time resolved CD measurements, as a function of CS^ESC concentration.

Concentration dependency of the initial folding rate was only investigated in samples prepared from stock 2. For measurements taken at pH 1.5, at 15°C and 21°C, we plotted the initial folding rate as a function of the CS^ESC concentration, resulting in **Figure 4.2 b**. As can be seen, there was no significant difference in initial folding rate between the two

temperatures. Therefore a single straight line was fitted through the data points belonging to both temperature data sets with the least squares method. The fit has an exponent of 1.8, suggesting that ribbon growth is a second order process, in contradiction to the first order kinetics suggested by the monoexponential fit to the time resolved CD₂₀₀ data.

4.3.4 Ribbons are attractive and tend to crosslink

When CS^ES^EC gels were formed we generally saw a small amount of water expelled from the gel after gel formation. A freshly formed gel drop of 206 mg and 20 g/l CS^ES^EC was allowed to mature under an equal amount of water for one week. After removal of the surrounding liquid, the gel drop appeared to have lost 1.5 % of its weight. When the 203 mg gel drop was left to dry in open air, it lost 60.7 mg in 14 hours. It was then submerged in the same solution that had been removed from before. After 14.5 hours it had regained 9.4 mg and after another 9.5 hours it had regained a total of 9.8 mg. The syneresis at gel formation, and the failure to swell more than 16% of the lost weight suggest that the ribbons in CS^ES^EC gels tend to form permanent crosslinks when brought into close contact. This implies attraction between the ribbons; hence we have attractive gels.

Microscopic observations in **Chapter 3** further support the attraction of CS^ES^EC ribbons to each other. In all microscopic (AFM, TEM and cryo-TEM) images obtained from 0.1 g/l ribbon dispersions, large areas of the substrate were found to be empty; when ribbons were found, they were usually clustered (**Fig 3.3 a** and **Fig. 3.4**) indicating either attraction and crosslinking or heterogeneous nucleation.

4.3.5 Dissolution of a 0.9 g/l CS^ES^EC gel by increasing pH

DLS was applied to detect Brownian motion in CS^ES^EC gels incubated for over six weeks at various pH. In the measurements, the intensity autocorrelation function g_2 at long times decays to unity for disintegrated gels in which the molecules can freely diffuse, but maintains a high value when Brownian motion is impeded by the gel state. The difference between g_2 at short and g_2 at very long times, Δg_2 , is therefore a measure for the gel-like behavior of the sample: a drop in the value of Δg_2 indicates the formation of a gel. We obtained the g_2 function and calculated Δg_2 for every sample and plotted it as a function of the pH (**Fig 4.3**). A 0.9 g/l CS^ES^EC gel at low pH, measured in this setup would normally give a Δg_2 value of 0.08. Increasing pH from pH 5.2 to pH 5.39 softens the gel as can be seen in **Figure 4.3** by the

gradual increase in Δg_2 . The low Δg_2 until pH 5.39 indicates gel-like behavior while the higher Δg_2 found above pH 5.45 is what it would be for a solution measured in this setup. This agrees with other observations. Up to pH 5.39, the presence of a gel at the bottom of the sample can be seen by the slightly turbid appearance of the gel, while at higher pH the entire sample appeared clear, homogeneous and liquid. According to these observations as well as the DLS measurements, a 0.9 g/l CS^ES^EC gel melts between pH 5.39 and pH 5.45.

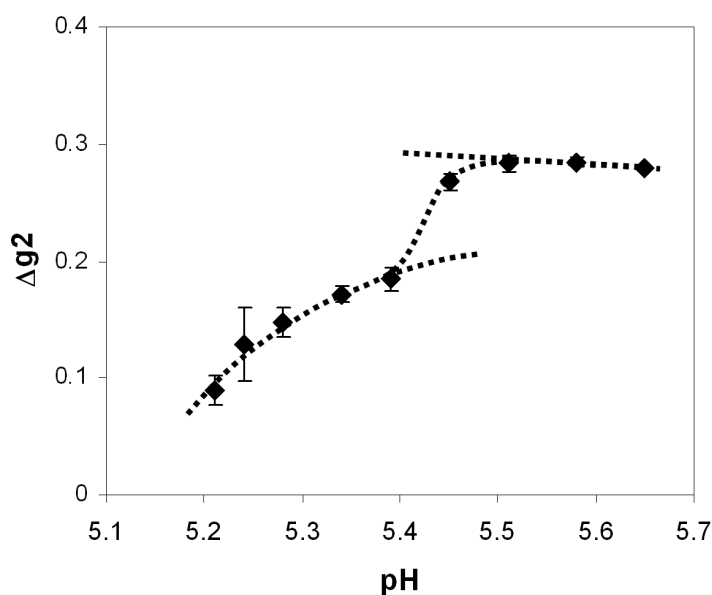


Figure 4.3 Δg_2 values for 0.9 g/l CS^ES^EC gels formed at pH 2 and equilibrated with 100 mM phosphate buffer at different pH, as a function of the final pH in the sample. Gel-like behavior is found for CS^ES^EC at pH < 5.39 and dissolved behavior for CS^ES^EC at pH > 5.45.

4.4 Discussion

In **Chapter 2** we showed that at low pH CS^ES^EC forms a gel, that dissolves again when the pH is increased to alkaline values. The present work (at an ambient temperature of 21°C), describes kinetic behavior of ribbon formation depending on pH and CS^ES^EC concentration, and also reports the pH at which the gel starts to dissolve.

Ribbon formation kinetics for most proteins have been reported to involve a nucleation step [1, 2], which manifests itself as a lag phase before ribbon growth is observed. According to CD data (**Fig. 4.1 a**) the ribbon formation reaction of CS^ES^EC does not display a lag phase, and the growth of the amount of ribbon is monoexponential, as in a first order process. This could be due to rate limiting monomolecular folding of the protein in solution. However, if ribbon growth of CS^ES^EC were a first order process with protein folding in solution as the rate

limiting step, the initial folding rate ($d[f]/dt$) would be directly proportional to $CS^E S^E C$ concentration, which we do not find (**Fig 4.2 b**). Instead, the initial folding rate is proportional to the $CS^E S^E C$ concentration as $d[f]/dt \sim [m]^{1.8}$, the exponent being closer to 2 than to 1. Thus a different explanation for the monoexponential curves is needed, like nuclei already being present in the sample, while nucleation is slow.

If the amount of nuclei spontaneously appearing per unit time is very small in comparison to the amount of nuclei (seed) that is already present, the concentration of growing ends of the ribbons will be approximately constant throughout the ribbon formation process. The probability that a free molecule meets and attaches to the end of a ribbon will therefore be simply proportional to the decreasing concentration of free molecules. Thus, if nucleation is slow and ribbon growth is relatively fast, one expects that seeded samples display monoexponential growth of the amount of ribbon. Equation 4.1 describes such monoexponential, irreversible growth of ribbons in a seeded solution

$$\frac{d[m]}{dt} = -k [e] [m] = \frac{-d[f]}{dt} \quad 4.1$$

in which $[m]$ (monomer) is the concentration of protein not in the ribbons, $[e]$ is the concentration of growing ends, $[f]$ is the concentration of protein present in the ribbons (fibril), t is time, and k is the reaction constant.

If a protein stock solution contains seed, this will also be the case in every sample prepared from this stock. As a function of concentration, $[e]$ will not be constant, but proportionate to the amount of protein used. In effect $[e]$ will be proportionate to $[m]$. Because of this, the initial folding rate $d[f]/dt$ as a function of total protein concentration $[m]_{(0)}$, will scale as $d[f]/dt \sim [m]_{(0)}^2$ for samples prepared from a seed containing stock solution.

Another indication for nucleation is the clustering of ribbons that we see with microscopy (**Fig 3.3 a** and **Fig.3.4**). Substrates dipped into a dilute ribbon suspension were often largely devoid of ribbons, but when ribbons were found, they were close together. Larger nuclei, possibly heterogeneous in nature, could give rise to the clustering of ribbons that we see with microscopy. Just after the acidification of the solution the nuclei will generate short growing ribbons. Because the ribbons are attractive and tend to crosslink when brought close together, the short, growing ribbons might stick and form the center of a future cluster. In this manner clustering would also occur, even if a nucleus would allow only two growing ends.

The formed ribbons are extremely stable at low pH, eventually resulting in the total conversion of free protein into μm long ribbons. We can convert the amplitudes of the CD signal at 200 nm into concentrations of folded protein and the derivative of the time resolved CD signal at $t=0$ into the initial rate of ribbon formation.

Accordingly, the initial rate of ribbon formation was studied as a function of pH. The data were gathered from two sets of samples, each set prepared from an individual stock solution. These two stock solutions were supposed to be equal to each other, yielding one set of data. To our surprise we found two different sets of data (**Fig 4.2 a**). Both sets of data show the same trend: a decrease of ribbon formation rate with increasing pH and complete inhibition of ribbon formation above pH 4.5. Still, the difference in conversion rate must be explained by some difference between the samples. The only difference between the samples was the protein stock solution used. The only difference between the protein stocks was that stock 1 was prepared freshly before measurements, while stock 2 was frozen and thawed several times before measurements. The difference in reaction speed between the two measured data sets is probably due to a different amount of seed for the two otherwise identical protein stock solutions. This difference may have been caused by the freezing and thawing of stock solution 2, which might have dissolved nuclei.

From **Figure 4.2 a** we may conclude that ribbon formation does not start before the pH is under 4.5 and according to **Figure 4.3**, a 0.9 g/l CS^ES^EC gel formed at pH 2 does not really dissolve until above pH 5.4. Thus, there seems to be hysteresis in the formation and the dissolution of the protein ribbons. At low pH the ribbons only grow at the ends, while at high pH the ribbon can fall apart at any location due to charge repulsion. It is conceivable however that just above pH 5.4 the ribbon does not fall apart at any location but dissolves from the ends. It is likely that the structure of a molecule in the middle of a ribbon is more stable than that of a molecule at an end. The structure of a folding or unfolding molecule at the end probably determines the energy barrier between the ribbon and free protein molecules, resulting in no ribbon formation or dissolution rate between pH 4.5 and pH 5.4.

4.5 Conclusion

Just as for natural protein fibrils, the CS^ES^EC molecules display nucleation and growth as the dominant mechanism for ribbon formation. The stock solutions appeared to contain nuclei from which fibrils in the prepared samples could grow. Nucleation seems to be slow

compared to ribbon growth, and so, for the seeded samples, we find a monoexponential increase of the amount of ribbon in time. The initial folding rate as a function of concentration suggests a second order reaction, also due to the stock solutions containing seed. The growing ribbons are at first short and are still free to move, to meet, and to stick to each other, possibly forming centers of future ribbon clusters. Otherwise clustering may be due to heterogeneous nuclei, from which many ribbons may start. Finally, a gel is formed in which the ribbons stick when they meet, resulting in an attractive gel which only shrinks but does not swell.

Acknowledgements

The authors thank Remco Fokink for his expertise on dynamic light scattering.

References

- 1 Murphy, R. M. (2007) Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins. *Biochimica Et Biophysica Acta-Biomembranes* **1768**, 1923-1934
- 2 Roberts, C. J. (2007) Non-native protein aggregation kinetics. *Biotechnology and Bioengineering* **98**, 927-938
- 3 Chen, P. (2005) Self-assembly of ionic-complementary peptides: a physicochemical viewpoint. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **261**, 3-24
- 4 Veerman, C., Rajagopal, K., Palla, C. S., Pochan, D. J., Schneider, J. P. and Furst, E. M. (2006) Gelation kinetics of beta-hairpin peptide hydrogel networks. *Macromolecules* **39**, 6608-6614
- 5 Yokoi, H., Kinoshita, T. and Zhang, S. G. (2005) Dynamic reassembly of peptide RADA16 nanofiber scaffold. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8414-8419
- 6 Zhang, S. G., Marini, D. M., Hwang, W. and Santoso, S. (2002) Design of nanostructured biological materials through self-assembly of peptides and proteins. *Current Opinion in Chemical Biology* **6**, 865-871
- 7 Haider, M., Megeed, Z. and Ghandehari, H. (2004) Genetically engineered polymers: status and prospects for controlled release. *Journal of Controlled Release* **95**, 1-26
- 8 Krejchi, M. T., Cooper, S. J., Deguchi, Y., Atkins, E. D. T., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Crystal structures of chain-folded antiparallel beta-sheet assemblies from sequence-designed periodic polypeptides. *Macromolecules* **30**, 5012-5024
- 9 Nagarsekar, A., Crissman, J., Crissman, M., Ferrari, F., Cappello, J. and Ghandehari, H. (2002) Genetic synthesis and characterization of pH- and temperature-sensitive silk-elastinlike protein block copolymers. *Journal of Biomedical Materials Research* **62**, 195-203
- 10 Smeenk, J. M., Otten, M. B. J., Thies, J., Tirrell, D. A., Stunnenberg, H. G. and van Hest, J. C. M. (2005) Controlled assembly of macromolecular beta-sheet fibrils. *Angewandte Chemie-International Edition* **44**, 1968-1971

5

Gelling kinetics and gel properties of silk-collagen-like block copolymers

A.A. Martens, J. van der Gucht, G. Eggink, F.A. de Wolf, M.A. Cohen Stuart

Abstract

Rheological data on monodisperse block copolymer hydrogels are rare because the amounts produced with various methods usually are not sufficient for material testing. With biotechnological methods, we produced enough material for a study of the mechanical properties of the protein block copolymer gels. To study the effect of block charge and block order, on the mechanical properties of self-assembling block copolymer hydrogels, we tested the physical behavior of $CS^E S^E C$, $S^E C C S^E$, and $CS^H S^H C$. Dynamic mechanical spectroscopy revealed differences in gelling kinetics and mechanical properties of the three different polymers. Remarkably, the S^E containing polymer gels displayed non linear elasticity comparable to that of the actin gels and other biological gels. Moreover, exceptionally high dynamic elasticity moduli, exceeding 40 kPa, were reached already at concentrations as low as 1.5 wt%, without any additional crosslinking agent. Such highly rigid yet dilute gels are rare and sought after. $CS^H S^H C$ gels were relatively weak and formed slowly. Additionally we studied the effect of temperature on the mechanical properties of a $CS^E S^E C$ gel.

5.1 Introduction

Dilute gels formed by percolation of a small amount of one component through a bulk of the other can contribute to the development of nanoporous materials for a wide range of applications varying from organic solar cells [1] to scaffolds for human tissue engineering [2].

Within tissue engineering, scaffold stiffness is an important physical factor in the response of many cell types [3]. The higher the scaffold rigidity, the better the cell adhesion and growth [3]. On the other hand low concentrations of gel-forming polymer are preferred, because this leads to higher porosity of the gel, which increases cellular mobility and nutrient transport [4]. This is also why the preferred mesostructure is such that no closed compartments are created.

Such dilute gels are usually formed by molecules that self-assemble in order to span a large volume with little material. There are some synthetic molecules that form such dilute hydrogels. Generally they are synthetic block copolymers [5], often containing synthesized polypeptides [4, 6, 7], and they are mostly produced using expensive, complex organic chemistry. Some natural proteins like actin fibrin and vimentin [8, 9] also form dilute gels, spanning the volume with self-assembled long filaments.

There are only few examples of rheological characterizations of gels made from large monodisperse block copolymers. For chemically synthesized monodisperse block copolymers there is no counterpart in literature. However there are a few examples concerning protein polymers [10, 11]. The sporadic occurrence of rheological data on large monodisperse block copolymers is probably due to the small amounts usually obtained with various synthetic methods. In the few cases that such rheological data is presented, biotechnology has been key in obtaining sufficient amounts of material for testing. The methods described in **Chapter 3**, similar to earlier methods [12, 13], have proven valuable as enabling technology for producing large monodisperse block copolymers, in quantities, sufficient for macroscopic investigations. The amounts produced (about 1 gram of each) allowed several rheological experiments to be conducted.

In this chapter we investigate what kind of macroscopic mechanical properties gels of self-assembled protein block copolymer ribbons possess, and how these properties are related to the gels mesoscopic structure. We investigated three different proteins: $CS^E S^E C$, $S^E CCS^E$ and $CS^H S^H C$. The difference between $CS^E S^E C$ and $S^E CCS^E$ is the block order. The difference

between $CS^E S^E C$ and $CS^H S^H C$ is that histidine replaces glutamic acid, wherever this occurs in the S^E block sequence. These may seem like minor details, but they have significant effects. Because the molecules are otherwise identical, we can further assess the influences of block order and block charge on gelling kinetics, microstructure and resulting (gel)material properties. Answering the questions above and studying the effect of block order and block charge, will enable us to tune macroscopic gel properties by changing these variables in possible future products. It might even help to develop the rational design of other protein-polymers containing blocks with different (structural) functions.

The investigated gels were formed by the self-assembly of a pH and charge responsive block, either S^E or S^H , as described previously in **Chapter 2, 3 and 4**. An analogue to the S^E block, with the same octapeptide repeat, but shorter, has previously been investigated as a separate species S^E , [14-16] or as part of a PEO- S^E -PEO hybrid triblock copolymer [17, 18]. The hybrid was very similar to our $CS^E S^E C$ product, being an S^E type block flanked by two hydrophilic blocks. It also self-assembled into well defined nano-sized ribbons [17, 18] similar to the ones that we found. However, rheological data on this S^E -analogue hybrid block copolymer is not available.

Here we successively discuss the behavior of the three different silk-collagen-like block copolymers mentioned above. For $CS^E S^E C$ at different concentrations, we discuss the gel formation kinetics and the material properties of the gel, all in relation to the nanostructure. Next, the unusual behavior of a $CS^E S^E C$ gel in response to changing temperature is shortly discussed. Finally, we discuss the kinetics and gel properties of $S^E CCS^E$ and of $CS^H S^H C$, and compare them to those of $CS^E S^E C$, drawing conclusions on the influence of block order and charged residue on the structure and material properties.

5.2 Materials and methods

5.2.1 Products and stock solutions

The products $CS^E S^E C$, $S^E CCS^E$ and $CS^H S^H C$ are all 802 amino acid long monodisperse block copolymers. The amino acid sequences and their production and purification have been described previously in **chapter 2**. Stock solutions of 10 g/l $CS^E S^E C$ and $S^E CCS^E$ were prepared by dissolving pure, freeze-dried material in 10 mM NaOH. A 10 g/l stock solution of $CS^H S^H C$ was prepared by dissolving pure freeze-dried material in 10 mM HCl.

5.2.2 Fibril concentration in time

One time resolved circular dichroism measurement at 200 nm of 0.2 g/l CS^ES^EC at pH 1.5, was done as described in **Chapter 4**. The monoexponential curve fitted to the CD data was converted to a curve that represents the concentration of fibrils in time (**Fig. 5.1 a**). This was done according to: $[f]_{(t)} = (E_{(t)} - E_0) * [p] / A$, in which $[f]$ is the fibril concentration, E is the (fitted) ellipticity, p is the total protein concentration and A is the amplitude of the curve fitted to the CD data.

5.2.3 Dynamic Light Scattering (DLS)

DLS was used to determine the time it takes to form a percolating network after sudden acidification (t_g). Upon gel formation, the autocorrelation function (g_2) does not decay fully between short and large correlation times as it would do for a solution [19]. The decay amplitude (Δg_2), which is the difference between g_2 at $\tau = 0$ and g_2 at large τ , drops when a percolating network (gel) is formed at t_g .

Samples of 0.1, 0.2 and 0.8 g/l CS^ES^EC were prepared at pH 1.5 from stock solution, water, and 1 M HCl. We measured at a fixed angle of 90° in a light scattering setup (ALV/DLS – 5000), equipped with an argon laser emitting vertically polarized light with a wavelength of 514.5 nm and a multiple τ digital correlator. Runs of 900 seconds were used, each run yielding an autocorrelation function g_2 and an amplitude Δg_2 , which was plotted as a function of time for each sample (**Fig. 5.1 b**).

5.2.4 Dynamic Mechanical Spectroscopy

Storage and loss moduli were determined for all measured gels using a rheometer (Physica MCR 300) with Couette CC17 geometry and a bob and gap size of 8.5 and 0.7 mm, respectively. All samples were prepared in similar fashion by mixing three components to a final amount of 6 ml: first an amount of stock solution of the protein, then the amount of water to arrive at the intended end concentrations and finally the acid or base needed to arrive at the intended pH. Transferring 4.2 ml to the rheometer was done within one minute after mixing. Each sample was topped off with 750ml low viscosity paraffin oil to prevent evaporation. **Table 5.1** contains the amounts used for every rheological sample discussed in this paper. The remaining 1.8 ml of sample was used to measure the sample pH. After loading a sample, gelling kinetics were followed by measuring the storage and loss modulus at 1 Hz

and 0.1 % deformation for several hours up to several days until the development of the moduli reached a plateau. **Figures 5.2 a-b** and **Figure 5.8 a** are examples of such measurements. This was followed by a frequency sweep from 0.1 Hz to 100 Hz and a strain sweep from 0.01% strain until well above the failure point of the gel (**Fig 5.3 a** and **Fig 5.8 b**).

	product concentration g/l	stock μ l	water μ l	acid or base μ l	pH
CS ^E S ^E C	1	600	4620	780 240 mM HCl	1.5
CS ^E S ^E C	4	2400	2748	852 250 mM HCl	1.5
CS ^E S ^E C	8	4800	252	948 250 mM HCl	1.5
CS ^E S ^E C	15				
S ^E CCS ^E	8	4800	252	948 250 mM HCl	1.5
CS ^H S ^H C	8	4800	963	237 1M NaOH	12.5

Table 5.1 Amounts and concentrations used for sample preparation for rheological measurements.

For 1 g/l CS^ES^EC, temperature dependency of the moduli was investigated by performing frequency sweeps at several temperatures going from 20°C to 80°C and back. For every frequency sweep, the value at 1. Hz was taken and plotted against the temperature at which the sweep was performed in **Figure 5.6**. After the frequency sweeps at different temperatures, the gel was strained until broken.

5.2.5 Temperature dependent CD spectrometry

CD measurements at 200 nm were recorded on a Jasco J-715 spectropolarimeter while varying the temperature. Both measured samples contained 0.1 g/l CS^ES^EC. The samples were composed from CS^ES^EC stock solution, water and HCl. and were transferred to a 0.1 mm quartz cuvette directly after preparation. After transferring the sample, pH was measured in the remaining liquid. The sample at pH 3 was prepared more than 16 h before the measurement and the sample at pH 1.5 was prepared more than 20 h before the measurement to ensure that the structural change after acidification was complete. The peltier controlled temperature was changed from 20°C to 98°C during which a measurement was taken every 0.1°C.

5.3 Results

5.3.1 Formation of a percolating network

For CS^ES^EC at pH 1.5, fibril concentration increases monoexponentially in time. **Figure 5.1 a** is an example for 0.2 g/l. This figure was obtained by converting the converting CD signal (ellipticity) to fibril concentration according to the materials and methods. The conversion factor in this case was 96 mdeg*1*g⁻¹, close to the 94 mdeg*1*g⁻¹ calculated for the average conversion factor in **Chapter 4**.

Fibril growth and free diffusion of fibrils continues until a certain moment (t_g) when a sample spanning network of fibrils is formed. This moment can be detected using dynamic light scattering (DLS). Upon network formation, (free) Brownian motion ceases, and correlations in intensity fluctuations of light scattered by a sample remain, even at long correlation times. Free diffusion of fibrils can be observed in DLS as a fully decaying autocorrelation function (g_2) whereas g_2 does not decay fully for sample spanning networks. The difference between g_2 at small and very large τ (Δg_2) is constant at first, when the sample is still liquid, but as soon as a percolating structure emerges, Δg_2 starts to decrease (**Fig. 5.1 b**).

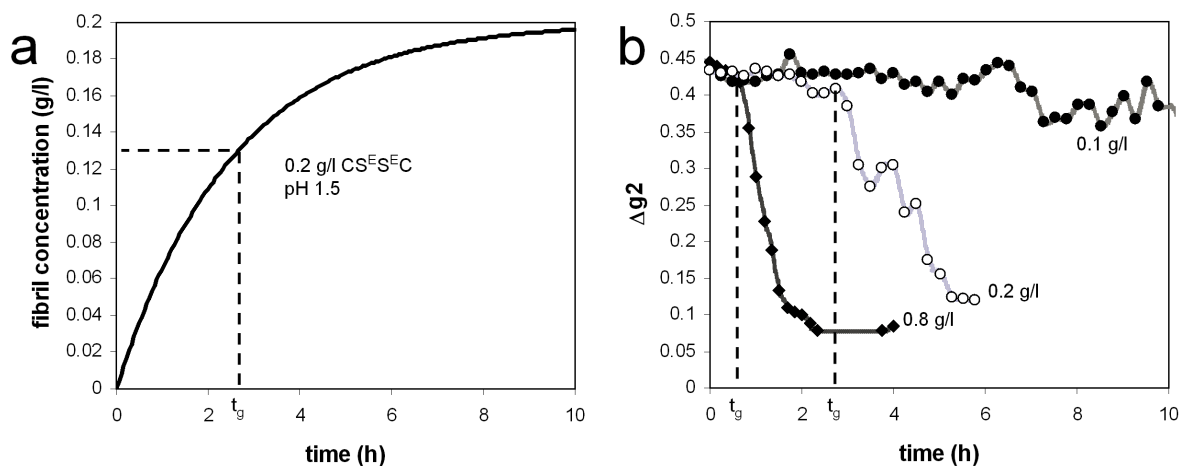


Figure 5.1 (a) Development of the fibril concentration in a 0.2 g/l CS^ES^EC sample just after bringing the pH to pH 1.5, based on time resolved DLS measurements at 200 nm. (b) From DLS measurements: Δg_2 , the difference between the value of the autocorrelation function at small and very large correlation times, as a function of time for various CS^ES^EC concentrations. At first, the value of Δg_2 is what it would be for a solution in this DLS setup, followed by a drop in Δg_2 , signifying the moment at which Brownian motion becomes impaired.

According to **Figure 5.1 b**, the time to network formation (t_g) for CS^ES^EC at pH 1.5 is concentration dependent. At 0.1 g/l no formation of a percolating network is observed using

DLS. For 0.2 g/l the time to network formation (t_g), visible as a drop in Δg_2 (**Fig. 5.1 b**), is 2.8 h. For 0.8 g/l, $t_g = 0.6$ h. The higher the $CS^{ESE}C$ concentration, the quicker a percolating network is formed.

5.3.2 Gel development

Gels of $CS^{ESE}C$, and $S^{ECCS}E$ were formed in a rheometer with a Couette configuration by changing the pH of the respective solutions. Solutions of $CS^{ESE}C$, and $S^{ECCS}E$ were acidified to pH 1.5 with HCl, and gel formation was followed by measuring the storage (G') and loss (G'') moduli at 1 Hz and 0.1% deformation. **Figure 5.2 a** is an example of such a gel formation measurement for 8 g/l $CS^{ESE}C$. At first, the solution had a very small storage modulus (black) and a stable loss modulus (gray). After an apparent lag time of 80 minutes they both started rising and the storage modulus surpassed the loss modulus indicating that the solution had become a gel. The gel reached 90% of its final G' in 27.4 h. In this example, the final storage modulus after 80 h was 6600 Pa.

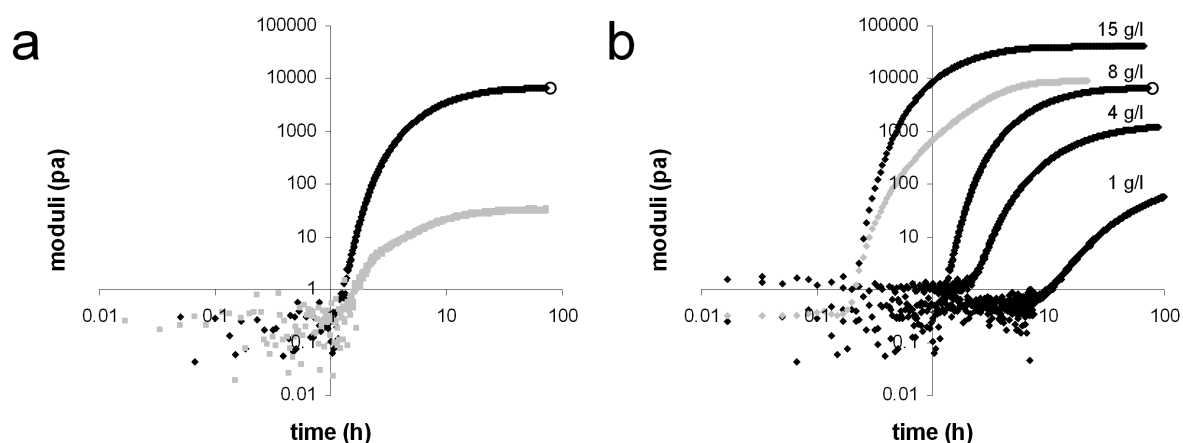


Figure 5.2 After shifting solutions to pH 1.5: (a) rheometric time courses of the build-up of the elastic modulus (black) and loss modulus (gray), for 8 g/l $CS^{ESE}C$, (b) rheometric time courses of the build-up of only the elastic moduli for: 1 g/l, 4 g/l, 8 g/l, and 15 g/l $CS^{ESE}C$ (black), as well as for 8 g/l $S^{ECCS}E$ (gray). Concentrations are written just above the corresponding data. For 8 g/l $CS^{ESE}C$, the same data was used in both figures a and b. This is indicated by the circle at the end of both curves. All measurements were taken in a Couette configuration at 1 Hz and 0.1% deformation.

In a series of samples with varying $CS^{ESE}C$ concentrations: 1, 4, 8 and 15 g/l of $CS^{ESE}C$ (**Fig. 5.2 b**), we find that the higher the concentration, the higher the final modulus and the shorter the lag time before the gel appears. In **Figure 5.2 b**, only the storage moduli are given for the different $CS^{ESE}C$ concentrations (black). Formation of an 8 g/l $S^{ECCS}E$ gel at pH 1.5 (**Fig. 5.2 b** gray), was also followed. At 8 g/l, $S^{ECCS}E$ gelled much faster than $CS^{ESE}C$. It had

a shorter lag time, as well as a higher dG'/dt just after the lag phase. The final modulus of 8 g/l S^ECCS^E was 8950 Pa, which is higher than the final 6600 Pa of 8 g/l CS^ES^EC .

For all samples, frequency and strain sweeps revealed that moduli were constant for frequencies from 0.1 Hz to 10 Hz and for strains from 0.01% to 1%. Thus, 1 Hz and 0.1% deformation were chosen as representative for this range of strains and frequencies, and were used for measuring a range of samples that could be compared.

5.3.3 Strain hardening

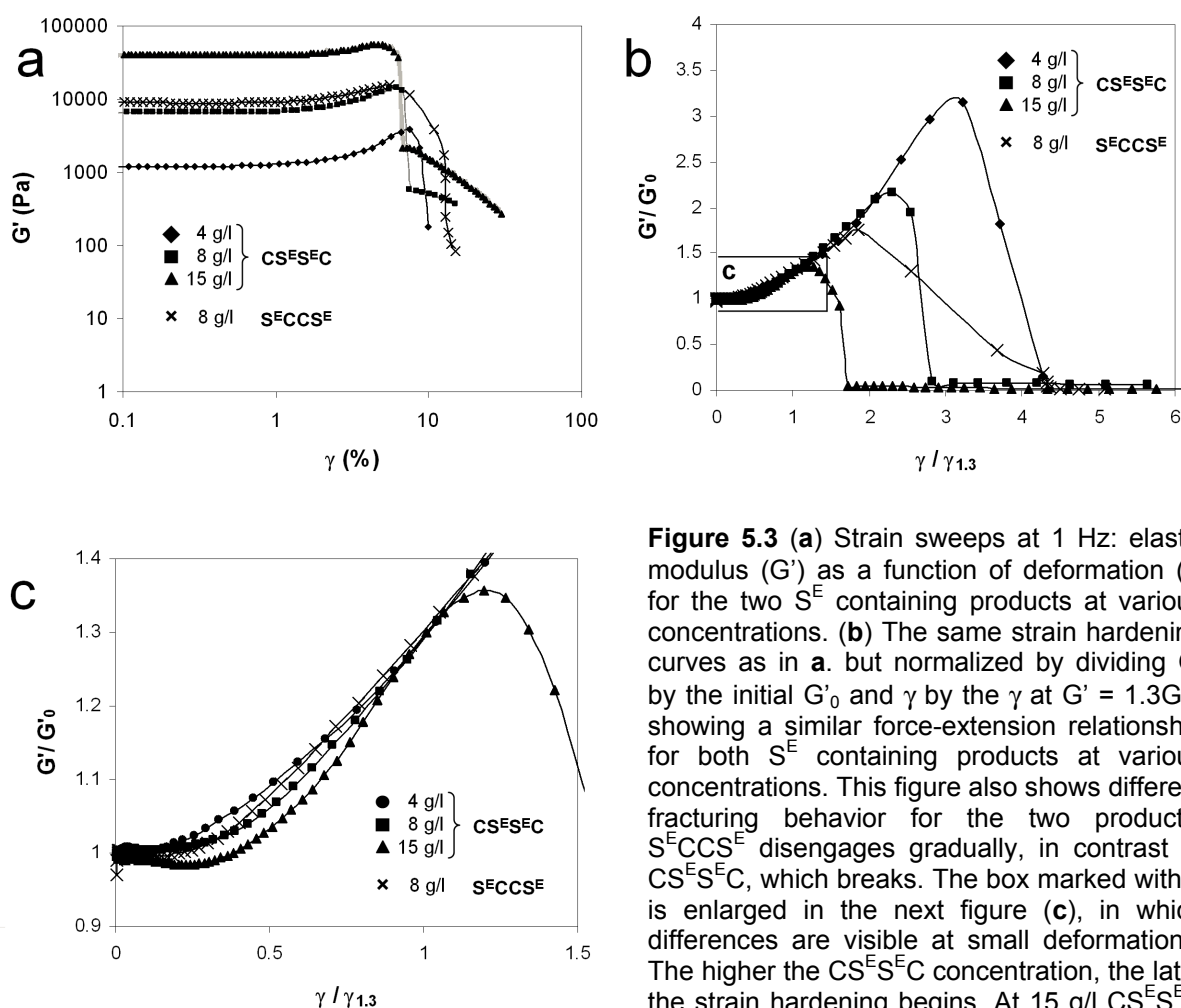


Figure 5.3 (a) Strain sweeps at 1 Hz: elastic modulus (G') as a function of deformation (γ) for the two S^E containing products at various concentrations. (b) The same strain hardening curves as in a. but normalized by dividing G' by the initial G'_0 and γ by the γ at $G' = 1.3G'_0$, showing a similar force-extension relationship for both S^E containing products at various concentrations. This figure also shows different fracturing behavior for the two products. S^ECCS^E disengages gradually, in contrast to CS^ES^EC , which breaks. The box marked with c is enlarged in the next figure (c), in which differences are visible at small deformations. The higher the CS^ES^EC concentration, the later the strain hardening begins. At 15 g/l CS^ES^EC there is even slight strain softening, before strain hardening sets in.

Gels that formed in the Couette cell were subjected to strain sweeps until failure (**Fig. 5.3 a**). Strain hardening was observed for the CS^ES^EC and S^ECCS^E gels. The storage moduli of the 8 g/l CS^ES^EC and 8 g/l S^ECCS^E gels were comparable: 6600, and 8950 Pa respectively.

The strain sweep curves $G'(\gamma)$ of the gels that displayed strain hardening were normalized to highlight their shape (**Fig. 5.3 b**). This was done by dividing G' by G'_0 and γ by $\gamma_{1.3}$ at which $G'/G'_{(0)} = 1.3$. The rising part of all curves coincides, pointing to a common strain behavior. An enlargement of a portion of **Figure 5.3 b** (**Fig. 5.3 c**) shows that there is only a slight difference in strain hardening behavior for the different concentrations of $CS^E S^E C$: the higher the concentration, the later the strain hardening begins. For 15 g/l the modulus even seems to decrease slightly before strain hardening sets in, but this is a very small effect.

5.3.4 Modulus as a function of $CS^E S^E C$ concentration

Measured storage moduli at 1Hz and 0.1% deformation were plotted as a function of $CS^E S^E C$ concentration (**Fig. 5.4**). For 1 g/l $CS^E S^E C$, a final modulus was estimated (165 Pa) based on monoexponential extrapolation of **Figure 5.2 b** using the MS Excel solver function. Not surprisingly, the storage modulus of gelled $CS^E S^E C$ at pH 1.5 was found to increase strongly with the protein concentration (closed rhombi). Fitting a power law curve to all four $CS^E S^E C$ data points gave: $G' = 147c^{2.0}$, with G' being the modulus (Pa) and c the protein weight concentration (g/l).

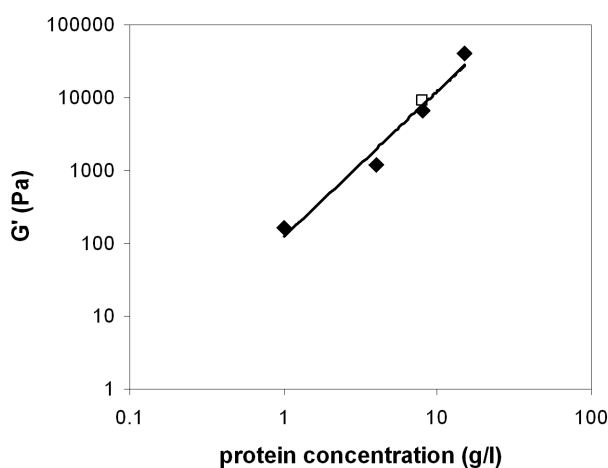


Figure 5.4 The elastic modulus (G') as a function of $CS^E S^E C$ concentration (◆). The fitted power law curve yielded an exponent of 2.0. The point for 8 g/l $S^E CCS^E$ (□) was also included for comparison with $CS^E S^E C$. All gels were measured at 1Hz and 0.1% deformation.

5.3.5 Strain at break as a function of $CS^E S^E C$ protein concentration

From the strain hardening curves, data points were derived for **Figure 5.5**, which displays maximal strain (γ_{max}) as a function of $CS^E S^E C$ concentration. The higher the $CS^E S^E C$

concentration, the more rigid the gel, and the lower the strain at which the gel fractured (closed rhombi). A power law fit gives $\gamma_{max} = 12.5c^{-0.36}$. The 8 g/l $S^E CCS^E$ gel broke at 5.48% strain (open square) which is slightly lower than the 6.10% strain at which the 8 g/l $CS^E S^E C$ gel broke.

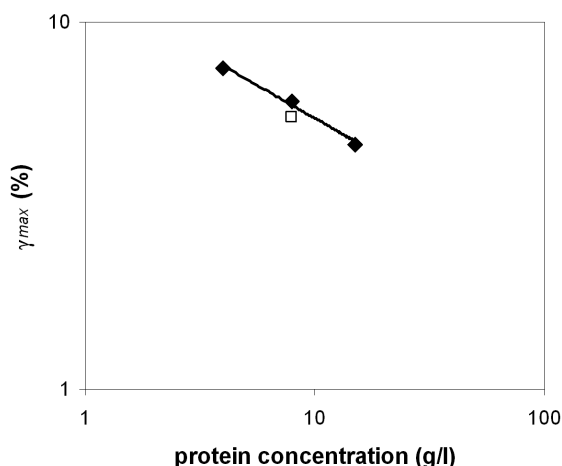


Figure 5.5 The deformation at which the gels fracture (γ_{max}) as a function of $CS^E S^E C$ concentration (◆) The fitted power law curve yielded an exponent of -0.36. The point for 8 g/l $S^E CCS^E$ (□) was also included for comparison with $CS^E S^E C$. All was measured at 1Hz.

5.3.6 Temperature effect on the modulus of a $CS^E S^E C$ gel

For a 1 g/l $CS^E S^E C$ gel in a Couette, the temperature was varied and the modulus was monitored (**Fig. 5.6**). Both the elastic modulus and the loss modulus (**Fig. 5.6 inset**) show pronounced hysteresis over the temperature cycle. When increasing the temperature from 20°C to 30°C, we first saw a slight decrease in modulus from 163 Pa to 155 Pa. Above 30°C the modulus increased almost twofold, reaching a maximum of 290 Pa at 50°C. Above 50°C the modulus rapidly declined to 37 Pa. Upon cooling the modulus never returned to its former values. Instead, it gradually decreased to 30 Pa. Later measurements showed no signs of gel recovery with time, and a strain sweep showed no strain hardening until the gel broke at 3% strain. Hence heating and cooling had permanently and drastically altered the gel.

5.3.7 Temperature effect on the CD signal of $CS^E S^E C$

Circular dichroism at 200 nm of acidic $CS^E S^E C$ solutions was followed as a function of temperature (**Fig. 5.7**). The black curve is a polynomial of the third degree and a least squares fit through the data points (black) gathered from a 0.1 g/l $CS^E S^E C$ solution at pH 3 at different

temperatures. The gray line was fitted similarly through data points gathered from a 0.1 g/l CS^ES^EC solution at pH 1.5, at different temperatures. The latter data points were not shown for clarity of the figure. Even though both data sets were noisy, the two fitted curves, at different pH were almost identical. From 20°C to 60°C the ellipticity at 200 nm increased slightly. Above 60°C it fell, probably due to the melting of the structure of the S^E block.

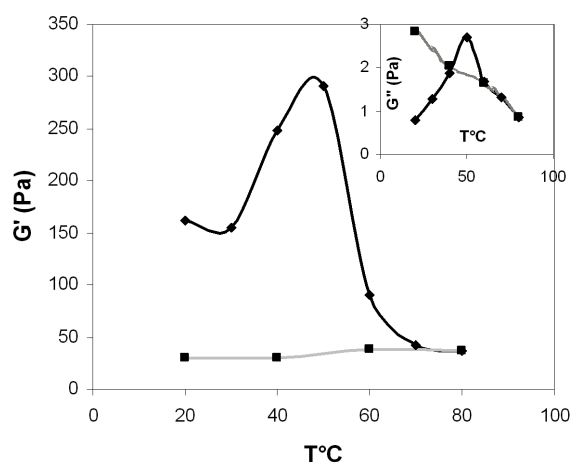


Figure 5.6 The elastic modulus (G') of a 1 g/l CS^ES^EC gel showing clear hysteresis in response to a temperature cycle with temperature increasing from 20°C to 80°C (♦) and decreasing 80°C to 20°C (■). Using the same symbols, the inset also shows an irreversible effect on the loss modulus. All was measured at 1Hz and 0.1% deformation.

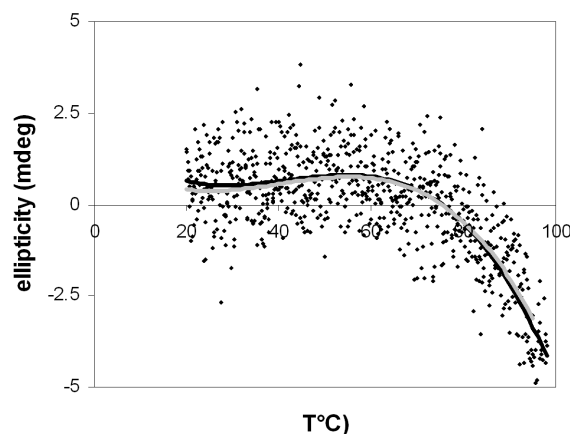


Figure 5.7 Datapoints are the CD signal at 200 nm for 0.1 g/l CS^ES^EC at pH 3 as a function of temperature. To these points, a polynomial of the third degree was fitted (black). Similarly, there was data at pH 1.5 (not shown) and the fitted curve (gray).

5.3.8 The effects of replacing glutamic acid with histidine in the S block

A gel of 8 g/l CS^HS^HC was formed in a rheometer with a Couette configuration by changing the pH to a value of 12.5 with NaOH. The gel formation was followed by measuring the storage (G') and loss (G'') moduli at 1 Hz and 0.1% deformation (**Fig. 5.8 a**). The gelling of 8 g/l CS^HS^HC was much slower than that of 8 g/l CS^ES^EC or 8 g/l S^ECCS^E and apparently had two stages of gel formation. The first stage started at a time of 1.7 h, when the storage modulus first appeared. After about 10 hours the modulus seemed to be constant, but at 24 hours the storage modulus began to increase again, revealing a second gelling stage. Frequency and strain sweeps revealed that moduli were constant for frequencies from 0.1 Hz to 10 Hz and for strains from 0.01% to 1%, confirming that, at 1 Hz and 0.1% deformation, measurements could be compared to the other samples.

The 8 g/l CS^HS^HC gel was subjected to a strain sweep until failure (**Fig 5.8 b**), and did not show any strain hardening. The G' of the 8g/l CS^HS^HC gel was an order of magnitude lower than the 8 g/l CS^ES^EC and 8 g/l S^ECCS^E gels, even though the protein concentration was the same.

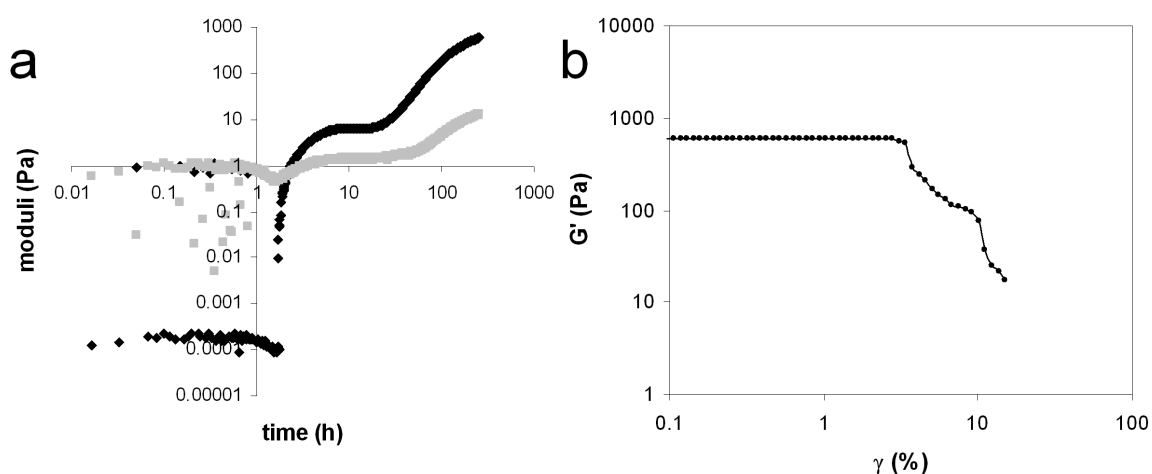


Figure 5.8 (a) Rheometric time courses of the build-up of the elastic modulus (black) and loss modulus (gray), for 8 g/l CS^HS^HC after shifting to pH 12.5, measured in a Couette configuration at 1Hz and 0.1% deformation. (b) Strain sweep at 1 Hz: elastic modulus (G') as a function of deformation (γ) for 8 g/l CS^HS^HC at pH 12.5.

5.4 Discussion

5.4.1 CS^ES^EC gelling kinetics

In a gelling sample, increasing fibril length and fibril concentration leads to network and therefore gel formation. As detected by rheology, network formation is apparently much slower than fibril formation. According to **Figure 5.1 a**, fibril formation is already 80 % complete after 4h, in a 0.2 g/l CS^ES^EC sample at pH 1.5, and according to **Chapter 4**, the fibril formation rate will only increase further with increasing CS^ES^EC concentration. This means that at 1 g/l, a higher concentration of which we did measure material properties in a rheometer (**Fig. 5.1 b**), fibril formation is probably completed well within 4 h, whereas at 1 g/l, we observe gelling in the rheometer only after 9 h (**Fig. 5.2 b**). So, it takes several hours after completion of the fibril formation, before the onset of the rise of moduli is detected by rheology.

Dynamic light scattering (DLS) enabled the determination of the time at which diffusion of the forming ribbons ceases. (**Fig. 5.1 b**). It appeared that, for most samples, diffusion ceases well before a significant increase in moduli is detected. It also appeared that, under the conditions of room temperature and pH1.5, a critical fibril concentration is needed for diffusion to cease. According to **Figure 5.1 b**, in 0.1 g/l CS^ES^EC diffusion never ceases, while for a slightly higher concentration of 0.2 g/l it does at $t_g = 2.8$ h. For the 0.2 g/l solution in **Figure 5.1 a**, at $t_g = 2.8$, the fibril concentration is 0.13 g/l. This is an estimate of the critical fibril concentration needed for diffusion to cease, and it is indeed higher than the 0.1 g/l in which diffusion does not cease (**Fig. 5.1 b**). For a 0.8 g/l solution, diffusion ceases even earlier, at $t_g = 0.6$ h.

Because no modulus can be detected at the moment that diffusion ceases, the network of fibrils is probably not yet crosslinked. Probably diffusion ceases when clusters of ribbons become large enough to touch each other, while the ribbons continue to grow, crossing from one cluster into the other, sterically locking the clusters in place. The increase in moduli that is detected rheometrically, long after t_g , and completion of the fibrils is probably caused by crosslinking between the ribbons of the immobilized microgels into a sample spanning network. Its formation apparently takes several hours.

So, it appears that the gels develop according to the following scenario. Upon acidification, fibrils start growing from nuclei (**Chapter 4**) into clusters. As the ribbon

concentration reaches 0.13 g/l, the clusters are so large that they inhibit each others diffusion and Brownian motion ceases. From then on, ribbons continue to grow, probably crossing from one cluster into the other. After immobilization of the clusters, it still takes up to several hours for the crosslinked, sample spanning network with a detectable modulus to develop.

Gelling kinetics were followed for various $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentrations, by monitoring the development of the storage and loss moduli in a Couette configuration (**Fig. 5.2 a,b**). In **Figure 5.2 b**, the development of only the storage moduli is depicted for $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ at concentrations varying from 1 g/l to 15 g/l. In this figure, the lag time before gelling can clearly be seen to decrease with increasing polymer concentration. This might be explained by t_g being reached much earlier and more ribbons crossing from one cluster into the other because more ribbon is formed after t_g . This increases the probability of crosslink formation between microgels. A similar explanation can be given for the rate at which the gel strength develops (dG'/dt). A larger amount of crosslinks forming per unit time, gives a larger increase in modulus per unit time, but more data would be needed for a quantitative analysis.

5.4.2 $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ gels consist of stiff crosslinked fibrils

To study the material properties of the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ gels after formation, the elastic modulus was measured as a function of strain for different concentrations (**Fig. 5.3 a**). The gels display strain hardening and a rather high modulus at low weight concentrations, (**Fig. 5.3 a**) which are two features, typical of a crosslinked network of semiflexible filaments [20]. Straight, apparently stiff and clustered fibrils were also observed in AFM and TEM images (**Fig. 3.3 a,c**) in **Chapter 3**.

A mechanism for elasticity, in networks of semiflexible filaments, that is entropic in origin, but can account for high moduli and strain hardening has been described for actin networks [20]. This model should be applicable to other semiflexible polymers, like ours, at intermediate concentrations. In this type of network, $a \ll \xi \leq l_p$ where a is the size of a monomer, ξ is the characteristic mesh size of the network, and l_p is the persistence length of a chain. As a consequence, the filaments do not form loops and knots, but they are sufficiently flexible to have significant thermal bending fluctuations [9]. For our network we assume permanent physical crosslinks with a characteristic crosslink distance l_c . The elastic properties arise from chains that are very nearly straight between crosslinks. Two effects contribute to the elasticity: bending of the chain and pulling out entropic fluctuations of the chain. “Internal

stretching” of the chain is neglected. This model [20] is valid for an isotropic, crosslinked gel of semiflexible chains. In this model, the modulus scales as

$$G' = \frac{l_p^2 k T}{\xi^2 l_c^3} \quad 5.1$$

Usually, for an entangled or crosslinked network of flexible polymers, ξ and l_c scale with the monomer concentration c as

$$\xi \sim l_c \sim c^{-1/2} \quad 5.2$$

resulting in

$$G' \sim c^{5/2} \quad 5.3$$

However, the concentration dependence of the crosslink length for semiflexible chains is different from that for flexible systems. For semiflexible chains, transverse fluctuations are greatly reduced over distances comparable to or smaller than the persistence length of the chain. In this model, the scaling of the crosslink length is assumed to be the same as that of the typical distance between binary collisions between chains in solution [20]. Thus l_c becomes larger than ξ for $\xi \ll l_p$. This leads to the crosslink length scaling with the concentration as

$$l_c \sim c^{-2/5} \quad 5.4$$

So finally the elastic modulus scales with the concentration as

$$G' \sim \xi^{-2} l_c^{-3} \sim c^{11/5} \quad 5.5$$

The measured plateau moduli depicted in **Figure 5.3 a** were plotted as a function of concentration in **Figure 5.4**, with the addition of one extra data point (1,100) estimated from extrapolating the modulus as a function of time for 1 g/l, depicted in **Figure 5.2 b**. G' in **Figure 5.4** scales with $c^{2.0}$. This is closer to the $c^{2.2}$ predicted by the model [20], than to the $c^{1.4}$ found for non-linked actin networks [21]. This is an indication that the model may apply to our gels and that they consist of crosslinked semiflexible filaments causing high moduli at low concentrations.

The fibrils themselves are anisotropic. We assume that the main load bearing structure of the fibril is a ribbon shaped core, made from the central $S^E S^E$ block, with according to **Chapter 3**, a cross-sectional height and width of 2.7 nm and 11.6 nm respectively. The crosslinks might be caused by hydrogen bonding between uncharged carboxylic acid groups

on the surface of the $S^E S^E$ block ribbons. If two ribbons were to cross each other perpendicularly, there would be $24 \times 24 = 576$ opposing pairs of uncharged carboxylic acid groups in the area of overlap between the ribbons, all capable of forming hydrogen bonds.

Because of the dimensions of the ribbon, it has a much higher persistence length in the width than in the height of the cross-section. Still, an estimate of the over all perceived persistence length of the fibrils can be made by combining the obtained mechanical data with structural data from **Chapter 3**.

First, we calculate the total length of fibrils (l_v) per volume of gel with a certain protein concentration. According to **Chapter 3 (Fig. 3.10)**, the β -roll constituting the hydrophobic core repeats every 0.95 nm of the ribbon. With a molar mass of 65750 g*mole⁻¹ (equivalent to 1.09×10^{-19} g molecule⁻¹) we have a linear density of 1.15×10^{-10} gm⁻¹ of fibril. From this linear density we can calculate the total length of fibril per unit volume (l_v) for different $CS^E S^E C$ concentrations. We can convert the result to a correlation length (ξ) by taking $\xi = l_v^{-0.5}$, this corresponds to the distance between two adjacent fibrils in a parallel arrangement, organized in squares perpendicularly to the fibril direction.

An estimate for the crosslink distance (l_c) may be done on the basis of l_v as follows. Assuming cubically arranged fibrils and crosslinks, the cubical volume (V) is spanned by three fibrils with an effective length of l_c each i.e. $V = l_c^3$. The fibril length per unit volume (l_v) should be equal to the total length of three fibrils, each with a length of l_c , divided by the volume that the three fibrils span: $l_v = 3 * l_c / V$. So, $l_v = 3 * l_c^{-2}$, and therefore the estimate for $l_c = 1.7 l_v^{-0.5}$. Probably l_c is larger, because the arrangement of the fibrils is most likely not cubical and also in cubical arrangement, fibrils may “miss” each other. The true length of one fibril ($l_c + \Delta l$) is slightly larger than l_c because the fibril is not fully stretched between the crosslinks, but because the ribbons are nearly straight, excess length of fibril between the crosslinks can be neglected.

By entering the estimated ξ , l_c , and the measured G' for various $CS^E S^E C$ concentrations into equation 5.1, the persistence length (l_p) can now be calculated. For the different concentrations 4, 8, and 15 g/l, persistence lengths of 14.8 14.5 and 16.4 μm were calculated respectively, with an average of 15 μm . Since l_c is likely to be larger than the estimate, and, according to equation 5.1 $l_p^2 \sim l_c^3$, the overall perceived persistence length must be at least 15 μm .

Interestingly, actin filaments have a very similar persistence length, namely $17\mu\text{m}$ [8]. Actin is a naturally occurring cytoskeletal protein with a high persistence length [8] and when crosslinked, it forms gels with exceptionally high modulus [8, 9, 22] much like ours. A similar persistence length is not surprising, because with a diameter of 7.5 nm, the dimensions of the cross-section of actin are close to those of our ribbon core ($2.8 * 14$ nm) even though the geometry of the cross-section is different. The main difference between our networks and actin networks is that the crosslink concentration in an actin gel can be controlled with the added amount of crosslinking, actin binding proteins (ABPs) [8], while our ribbons crosslink spontaneously. Our gels also showed some similarity to gels of a designed protein polymer denoted as poly-EAK9 [10]. The common factor between our gels, actin gels and poly-EAK9 gels is that they all consist of crosslinked semiflexible filaments.

5.4.3 Strain hardening and gel rupture

Above, we discussed the concentration dependency of the modulus, but **Figure 5.3 a** also showed strain hardening, the second trait of crosslinked networks of semiflexible chains, common for several different biological filaments, including actin [9]. To be able to compare the force-extension behavior for the different gels, the curves of the gels displaying strain hardening in **Figure 5.3 a** were normalized [9]. After normalization, $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ gels at different concentrations can, be seen to exhibit a similar force extension mechanism.

At small strains, the deformation of the gel is nonaffine [23]. This means that the deformation that is imposed on the gel does not propagate uniformly throughout the gel, i.e. some parts of the gel are more deformed than others. A tell tale sign of this is that at higher densities of filaments, the stiffness first decreases with strain at small strain levels as can be seen in **Figure 5.2 c**. This is caused by buckling of a fraction of filaments, loaded primarily in compression [23]. Since our ribbons are anisotropic, buckling will preferably occur in one direction, with the smallest ribbon dimension having the smallest radius of curvature. At higher deformations, an increasing number of filaments will stretch. Once stretching has set in, the deformation becomes more and more affine [23], i.e. the deformation is more and more homogeneously distributed in the gel. This results in similar normalized force extension curves at larger deformations.

Because of the normalization, all curves (**Fig. 5.3 b**) will by default pass through the points (0,1) and (1, 1.3), divergences would mainly become apparent above (1,1.3). Yet, all

curves continued running parallel to each other beyond (1, 1.3) indicating that the mechanism for strain hardening is the same for the various concentrations of CS^ES^EC. The gels display strain hardening in a way similar to that found in crosslinked semiflexible biological protein filaments such as actin [8, 9].

The deformation and the strain hardening continues until the gel ruptures. The maximum strain (γ_{\max}) at which the gel ruptures, as a function of CS^ES^EC concentration, also behaves as for crosslinked networks of semiflexible filaments. The line fitted to the data represented in **Figure 5.5** has an exponent of -0.36, which, seen the few data points, is indeed comparable to the theoretical value of -0.5 for such networks [8].

5.4.4 Effect of temperature on the moduli

Temperature has a rather unusual effect on the moduli of a 1 g/l CS^ES^EC gel. We changed the temperature from 20°C to 80°C and back, while monitoring the storage and loss modulus (**Fig. 5.6**).

Raising the temperature from 20°C decreased the modulus slightly, until 30°C, after which the modulus started to increase, reaching a maximum, where G' is twice as high, at about 50°C. The twofold increase over only 10°C is much more than one would expect for a rubber. According to equation 5.1, a rise in G' can be caused by an increase in persistence length (l_p) or a decrease in correlation length (ξ) or crosslink distance (l_c). It is not obvious which of these it is. There is no apparent reason why l_p would change much with temperature and the latter two, (ξ and l_c) are mainly concentration dependent, so it seems unlikely that they would change. Let us discuss the three cases.

The least likely explanation would be that ξ changes. It is only dependent on the length per unit volume of fibril. This could only change if the β -roll in the ribbon core would change into another conformation that contributes more to the fibril length than the β -roll. Such a conformational change would probably weaken the fibrils and the gel instead of strengthen it.

A possible, but not very likely explanation is that l_p increases with temperature between 40°C and 50°C. A large increase in modulus with temperature can be observed in gels containing elements with LCST behavior. However, no LCST behavior was observed in solutions of pure C (block) molecules at low concentrations. Still, on the ribbon, the C blocks are concentrated into a bottle brush-like structure up to values of locally more than 90 g/l. Under these conditions, the C block might display different behavior than at a low

concentration, as a single component in solution. Generally, solubility of proteins with high proline content, like the C block, decreases with temperature. At such high concentrations as in the corona, this might lead to a slight stiffening of the corona, contributing to the persistence length of the filaments. Since the modulus scales with the persistence length as $G' \sim l_p^2$, The increase of l_p by a factor 1.4 would already be sufficient to increase G' by a factor of 2.

A more simple explanation is a decreasing l_c due to formation of more crosslinks upon heating. The crosslink density is lower for semiflexible filaments than for flexible ones, because they are straight and, when trapped in a network, less likely to meet to form new crosslinks. With increased temperature, the increased thermal fluctuations of the filaments may cause previously unconnected filaments to touch and to crosslink, thus decreasing l_c and increasing the modulus. To further investigate this, thermal reversibility over only the 20°C to 50°C range should be tested. If the higher modulus is retained upon cooling from 50°C to 20°C, the new crosslinks are irreversible and crosslinking will be by far the most likely explanation.

Above 50 °C the modulus dropped. Parallel to this, the CD signal at 200 nm (**Fig. 5.7**) started to change above 60°C in favor of random structures. The decrease in modulus above 50°C may therefore be due to weakening of the crosslinks (probably hydrogen bonds), followed by structural change of the S^E block above 60°C which shows up in CD measurements (**Fig. 5.7**). At 80°C a frequency sweep still gives a constant G' down to 0.1 Hz, indicating that there still are long lived (permanent) network connections on this timescale.

When lowering the temperature from 80°C to 20 °C the storage modulus did not recover. It even decreased slightly and stayed stable for more than 4 h. If recovery would have occurred on a similar timescale as the initial gel formation, recovery should have become apparent within these 4 h. As can be seen in the inset of **Figure 5.6** the same temperature cycle also produced an irreversible effect on the loss modulus. After the temperature cycle, the gel did not longer show strain hardening, suggesting that there were no longer any crosslinks. As seen from the CD measurement (**Fig. 5.7**), heating led to conformational changes of the molecules. This in turn could have degraded crosslinks or broken the filaments which failed to reconnect upon cooling. We conclude that both the CD signal at 200 nm and G' should be monitored while increasing and decreasing the temperature. To investigate reversibility of the structures over parts of the temperature cycle, smaller intervals could be

taken. For example: between 50°C and 60°C we see a decrease in G' , but no conformational change according to the CD measurements. If this decrease in G' would be only due to the degradation of crosslinks and the crosslinks could reform upon cooling, G' would be reversible along this temperature interval.

5.4.5 S^ECCS^E gelling kinetics and gel properties

After acidification, both 8 g/l $CS^ES^E C$ and 8 g/l S^ECCS^E had at first very small storage and stable loss moduli (**Fig. 5.2 b**). After a much shorter lag time, the modulus of S^ECCS^E started rising much quicker than that of $CS^ES^E C$ (**Fig. 5.2 b**), i.e. S^ECCS^E is a much faster gelator than $CS^ES^E C$. For the different gelling curves we discuss two tentative explanations. Firstly, fibril formation might be quicker. The zipper-like structure of S^ECCS^E molecules, as discussed in **Chapter 3**, might have “split ends” leading to a growing number of growing ends in time, thus speeding up the gelling. Two outer S^E blocks are probably more exposed and prone to aggregation than one central S^ES^E block, leading to both faster fibril nucleation and fibril growth. Time resolved CD measurements on aggregating S^ECCS^E would be needed to test these hypotheses. Secondly, crosslink formation could be quicker. This may be because of the different corona conformation, (loops instead of ends) or because of increased crosslinking between different ribbons through shared middle blocks. Whatever mechanism is operating, we observe clear differences between the two samples, illustrating that even though the amino acid composition of the molecules and the primary sequence of the separate blocks is exactly the same, differences in gelling kinetics should be expected and might even be tailored by changing block order.

Once the 8 g/l S^ECCS^E gel was formed, it displayed the same typical features of a crosslinked network of semiflexible filaments as the $CS^ES^E C$ gels (**Fig. 5.3 a**). At 8900 Pa, it had a slightly higher modulus than the 6600 Pa of the 8 g/l $CS^ES^E C$ gel (**Fig. 5.4**). Consistent with a slightly higher modulus, it also displays a slightly lower γ_{max} (**Fig. 5.5**) Both might have been caused by more crosslinking. Still they are within the same order of magnitude as for $CS^ES^E C$.

Also after normalizing (**Fig. 5.3 b**), the S^ECCS^E gel appeared to have the same force extension relationship as the $CS^ES^E C$ gels. However upon failure, $CS^ES^E C$ gels broke abruptly while S^ECCS^E disengaged more gradually. Possibly, part of the S^ECCS^E molecules have their S^E blocks in different fibril cores, maintaining a weaker structure of core segments

crosslinked by the hydrophilic CC middle blocks, after the long fibril cores have been broken into smaller segments.

5.4.6 CS^HS^HC gelling kinetics and gel properties

Incorporating histidine, as the charge carrier on the S block had a huge impact on the behavior of the block copolymer CS^HS^HC, otherwise identical to CS^ES^EC. Not only did gelling occur at high pH instead of at low pH (this was expected) but also gelling kinetics (**Fig. 5.8 a**) and gel properties (**Fig. 5.8 b**) were different. Firstly, the gelling kinetics were much slower for CS^HS^HC than for CS^ES^EC and showed two distinct stages of gel hardening. Possibly the initially formed fibrils are stiffer than the CS^ES^EC fibrils, resulting in a low initial modulus (6.4 Pa), but also in slower crosslink formation, which starts only after 25 hours. However, the CS^HS^HC gel showed no signs of strain hardening (**Fig. 5.8 b**), indicating no permanent crosslinks.

Alternatively, the first gelling may be due to random aggregation after charge neutralization of the histidine residues at high pH. The 25 h that it takes for the second change to begin suggests a cooperative process like a conformational change which will only run after nucleation. Maybe a random network is formed first, after which some slow fibril formation takes place. It would be interesting to see if CD data could support two periods with two different dominant conformations.

The final modulus for 8 g/l CS^HS^HC was a decade lower than for 8 g/l CS^ES^EC. Moreover, SAXS measurements (**Chapter3**) also showed a possible combination of random or globular structures with ribbons. All these data show that changing the glutamic acid residues to histidine residues in the primary amino acid sequence not only inverts the pH response but also affects the preferred secondary and tertiary structures that appear upon charge neutralization of the molecule. The CS^HS^HC gel behavior is reminiscent of the thermocycled CS^ES^EC gel. Both CD spectroscopy and SAXS could be used to find any structural similarity on a molecular level.

5.5 Conclusion

At low pH, CS^ES^EC and S^ECCS^E produced gels with exceptionally high elastic moduli at low concentrations. This could be explained by networks of crosslinked semiflexible filaments. More exposed S^E blocks in S^ECCS^E may for several reasons cause S^ECCS^E to form ribbons and gels faster than CS^ES^EC does.

S^ECCS^E may have formed more crosslinks than CS^ES^EC , resulting in S^ECCS^E having a higher G' and lower γ_{max} than CS^ES^EC at the same wt%. Both molecules probably have the same crosslinking mechanism, but the CC middle blocks of S^ECCS^E molecules, having their two S^E blocks sitting in different fibril cores form additional crosslinks.

The unusual increase of the modulus of CS^ES^EC at increasing temperature might be explained by extra crosslinking due to extra thermal fluctuations of the fibrils. A further increase of temperature, apparently first causes crosslinks or fibrils to dissolve above 50°C and then melt the S^E block above 60°C probably causing massive breakdown or modification of filaments. After cooling the filaments apparently cannot reconnect to reform the crosslinked network.

Gel formation of CS^HS^HC was slow and occurred in two separate gel hardening stages. This indicates that after aggregation, cooperative conformational change led to gel hardening. replacing glutamic acid with histidine in the charged S block did not only invert the response to pH, but apparently also changed the filament and gel structure itself. The modulus of the CS^HS^HC gel was ten times lower than that of the CS^ES^EC gel. In accordance with the low modulus, the lack of strain hardening indicates that the CS^HS^HC gel is not crosslinked.

Of all products discussed in this chapter, S^ECCS^E would be the most useful for building tissue engineering scaffolds, because at 8 g/l it gels within minutes, in contrast to the 1.3 h that it takes for CS^ES^EC to start gelling at this (low) concentration. However, to this end, the structure must be produced at a biologically suitable, higher pH, by compensating the S^E block charge with an oppositely charged polyelectrolyte, which might lead to different mesostructures.

Acknowledgements

The authors thank Joris Sprakel from the laboratory of physical chemistry and colloid science in Wageningen for sharing his knowledge and giving instructions on dynamic mechanical spectroscopy.

References

- 1 Heiser, T., Adamopoulos, G., Brinkmann, M., Giovanella, U., Ould-Saad, S., Brochon, C., van de Wetering, K. and Hadziioannou, G. (2006) Nanostructure of self-assembled rod-coil block copolymer films for photovoltaic applications. *Thin Solid Films* **511**, 219-223
- 2 Hamada, K., Hirose, M., Yamashita, T. and Ohgushil, H. (2008) Spatial distribution of mineralized bone matrix produced by marrow mesenchymal stem cells in self-assembling peptide hydrogel scaffold. *Journal of Biomedical Materials Research Part A* **84A**, 128-136

- 3 Lv, Q., Hu, K., Feng, Q. and Cui, F. (2008) Fibroin/collagen hybrid hydrogels with crosslinking method: Preparation, properties, and cytocompatibility. *Journal of Biomedical Materials Research Part A* **84A**, 198-207
- 4 Pakstis, L. M., Ozbas, B., Hales, K. D., Nowak, A. P., Deming, T. J. and Pochan, D. (2004) Effect of chemistry and morphology on the biofunctionality of self-assembling diblock copolypeptide hydrogels. *Biomacromolecules* **5**, 312-318
- 5 Tsitsilianis, C., Iliopoulos, I. and Ducouret, G. (2000) An associative polyelectrolyte end-capped with short polystyrene chains. Synthesis and rheological behavior. *Macromolecules* **33**, 2936-2943
- 6 Nowak, A. P., Breedveld, V., Pine, D. J. and Deming, T. J. (2003) Unusual salt stability in highly charged diblock co-polypeptide hydrogels. *Journal of the American Chemical Society* **125**, 15666-15670
- 7 Morihara, Y., Ogata, S., Kamitakahara, M., Ohtsuki, C. and Tanihara, M. (2005) Thermosensitive gel formation of novel polypeptides containing a collagen-derived pro-hyp-gly sequence and an elastin-derived val-pro-uy-val-gly sequence. *Journal of Polymer Science Part a-Polymer Chemistry* **43**, 6048-6056
- 8 Gardel, M. L., Shin, J. H., MacKintosh, F. C., Mahadevan, L., Matsudaira, P. and Weitz, D. A. (2004) Elastic Behavior of cross-linked and bundled actin networks. *Science* **304**, 1301-1305
- 9 Storm, C., Pastore, J. J., MacKintosh, F. C., Lubensky, T. C. and Janmey, P. A. (2005) Nonlinear elasticity in biological gels. *Nature* **435**, 191-194
- 10 Goeden-Wood, N. L., Conticello, V. P., Muller, S. J. and Keasling, J. D. (2002) Improved assembly of multimeric genes for the biosynthetic production of protein polymers. *Biomacromolecules* **3**, 874-879
- 11 Nagapudi, K., Brinkman, W. T., Thomas, B. S., Park, J. O., Srinivasarao, M., Wright, E., Conticello, V. P. and Chaikof, E. L. (2005) Viscoelastic and mechanical behavior of recombinant protein elastomers. *Biomaterials* **26**, 4695-4706
- 12 Werten, M. W. T., Van den Bosch, T. J., Wind, R. D., Mooibroek, H. and De Wolf, F. A. (1999) High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**, 1087-1096
- 13 Werten, M. W. T., Wisselink, W. H., van den Bosch, T. J. J., de Bruin, E. C. and de Wolf, F. A. (2001) Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Engineering* **14**, 447-454
- 14 Cantor, E. J., Atkins, E. D. T., Cooper, S. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Effects of amino acid side-chain volume on chain packing in genetically engineered periodic polypeptides. *Journal Of Biochemistry* **122**, 217-225
- 15 Krejchi, M. T., Atkins, E. D. T., Waddon, A. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1994) Chemical Sequence Control Of Beta-Sheet Assembly In Macromolecular Crystals Of Periodic Polypeptides. *Science* **265**, 1427-1432
- 16 Krejchi, M. T., Cooper, S. J., Deguchi, Y., Atkins, E. D. T., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Crystal structures of chain-folded antiparallel beta-sheet assemblies from sequence-designed periodic polypeptides. *Macromolecules* **30**, 5012-5024
- 17 Smeenk, J. M., Otten, M. B. J., Thies, J., Tirrell, D. A., Stunnenberg, H. G. and van Hest, J. C. M. (2005) Controlled assembly of macromolecular beta-sheet fibrils. *Angewandte Chemie-International Edition* **44**, 1968-1971
- 18 Smeenk, J. M., Schon, P., Otten, M. B. J., Speller, S., Stunnenberg, H. G. and van Hest, J. C. M. (2006) Fibril formation by triblock copolymers of silklike beta-sheet polypeptides and poly(ethylene glycol). *Macromolecules* **39**, 2989-2997
- 19 Romer, S., Scheffold, F. and Schurtenberger, P. (2000) Sol-gel transition of concentrated colloidal suspensions. *Physical Review Letters* **85**, 4980-4983
- 20 Mackintosh, F. C., Kas, J. and Janmey, P. A. (1995) Elasticity of Semiflexible Biopolymer Networks. *Physical Review Letters* **75**, 4425-4428
- 21 Hinner, B., Tempel, M., Sackmann, E., Kroy, K. and Frey, E. (1998) Entanglement, elasticity, and viscous relaxation of actin solutions. *Physical Review Letters* **81**, 2614-2617

- 22 Janmey, P. A., Hvidt, S., Lamb, J. and Stossel, T. P. (1990) Resemblance of Actin-Binding Protein Actin Gels to Covalently Cross-Linked Networks. *Nature* **345**, 89-92
- 23 Onck, P. R., Koeman, T., van Dillen, T. and van der Giessen, E. (2005) Alternative explanation of stiffening in cross-linked semiflexible networks. *Physical Review Letters* **95**

6

General discussion and conclusion

A.A. Martens

6.1 General discussion

We set out to make new polymeric materials that are superior to existing materials. More specifically, this research was intended to lay the basis for the development of fibers that could measure up to the strength of Dynema® [1] and Twaron® [2] in the direction parallel to the fiber, but be superior in strength in directions perpendicular to the fiber (from here on called parallel and transversal direction). This chapter describes in detail the initial design considerations for molecules that could potentially constitute such a fiber. It also discusses the consequences that the final design had on the self- and co-assembled structures and gels described in this thesis. These gels and structures appeared to be potentially suitable for other applications while still offering a way forward in creating strong fibers.

The “secret” of the super strong fibers is that during polymer processing, long polymer chains are aligned parallel to each other. The direction of alignment is also the direction of largest strength. Such is the case for Dyneema® [1], which is high molecular weight polyethylene in which the polymers are aligned in a process now called gel spinning [3]. In the parallel direction it is the strongest fiber produced until now (stronger than steel) because pulling on the stretched carbon chains is like pulling on all carbon-carbon atom bonds simultaneously. Transversally however, there are no other than Vander Waals forces keeping the polymers together, basically making the fiber weak in this direction [4]. In the case of Twaron® [2], which (like protein) is a polyamide, hydrogen bonding in the transversal direction, makes the fiber transversally slightly stronger. So in order to have even larger transversal strength, we set out to make large molecules (**Fig. 6.1**) that have polymeric blocks with a strong attraction to each other, forming crystalline physical crosslinks between the molecules in the transversal direction (**Fig. 6.1**), but that also contain flexible polymeric blocks that can be stretched in the parallel direction (**Fig. 6.1**), to produce a fiber of aligned molecules. Moreover, for efficient packing, the size of the “sticky” blocks should all be the same and the size of the flexible blocks should all be the same, constituting identical molecules (**Fig. 6.1**). Self-assembly is thought to benefit from monodisperse block size [5]. In this way, by neat self-assembly, the interaction per “sticky block” will be maximized, because there are no uncovered longer blocks “sticking out” of the crystalline domain or shorter blocks creating a weak notch in the crystalline domain. This will benefit the strength especially in the transversal direction. If all flexible blocks have the same length and also the

same distance to bridge between the crystalline domains, we will, when pulling on the fiber, be pulling equally on all individual flexible blocks, which will benefit the strength in the parallel direction. All in all, using large monodisperse block copolymers with monodisperse blocks would reduce the number of defects in the material and improve the strength in both fiber directions.

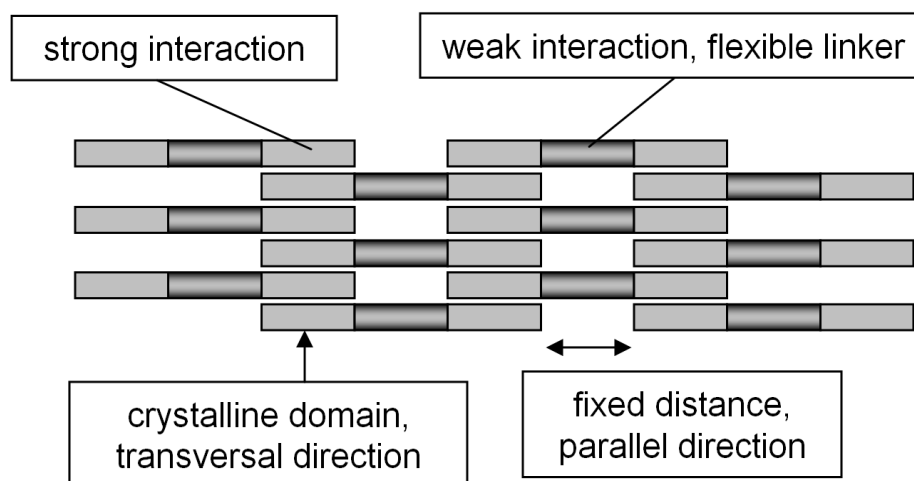


Figure 6.1 Cartoon of block copolymers containing attractive blocks that self-assemble into crystalline domains and flexible blocks that can link the crystalline domains.

To test this idea, we needed large, monodisperse block copolymers, with monodisperse blocks, some with strong and some with weak interaction to each other. For the strong interaction we chose the Coulombic interaction between positively and negatively charged polymers which is far stronger than hydrogen bonding [6]. However if positively and negatively charged blocks were used within one polymer, the polymer would fold back on itself making mainly loops and not so much networks. Therefore it would be better to separately produce two block copolymers, one positively charged and the other negatively charged, that both are soluble as separate components (**Fig. 6.2 a,b**). When mixed, their charged blocks are supposed to form crystalline domains similar to those in spider silk [7] but then with additional charge complexes serving as the main physical crosslinking force between the molecules (**Fig. 6.2 c**). One could also say that, in such a structure, the flexible blocks are covalently attached and flexible crosslinks between the crystalline domains, thus forming a molecular network of oppositely charged block copolymers (**Fig. 6.2 c**). To summarize the requirements of the molecules, we desire: at least two, long (hundreds of

monomers), oppositely charged polymers containing charged blocks of exactly equal size and flexible blocks of exactly equal size.

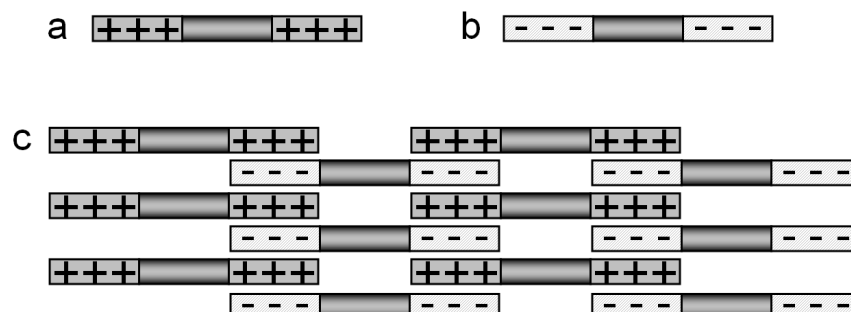


Figure 6.2 Cartoon of (a) positively charged block copolymer (b), negatively charged block copolymer, which can both be produced separately as soluble components and (c) a molecular network that these two components might form if the charged blocks crystallize with each other.

Unfortunately such molecules were not available. Since classical polymer chemistry could not provide them, we had to turn to other means of obtaining such large and monodisperse block copolymers. One way of doing this is by exploiting the natural protein production machinery of the living cell [8, 9]. It works as follows: just as a natural gene (DNA) can be transcribed into RNA, which then can be translated by ribosomes that synthesize the natural protein encoded in the original DNA template, we can insert a synthetic gene into a host organism that will be transcribed and translated into our synthetic polypeptide (protein block copolymer). We chose a production organism called *Pichia pastoris* GS115 [10], in combination with the pPIC9 expression cassette [11]. This combination is well known for its generally high production yield [11]. Importantly, *P.pastoris* also has a good track record of producing repetitive amino acid sequences [9, 12-14], which a polymer like ours would be.

Since we chose the production mechanism “protein expression in *P.pastoris* of a synthetic gene encoding a monodisperse block copolymer”, fulfilling the requirements that the product should be monodisperse and sequential, we were limited in our polymer design to the 20 natural amino acids, which still leaves a lot of choice in different monomers. Having 20 different chiral monomers (amino acids) with hydrophobic, hydrophilic, acidic, basic, aromatic and reactive side groups, all displaying different physiochemical properties and conformational preferences and being able to polymerize them in any desired sequence is still highly versatile and forms hardly a limitation on the broad range of potential products, as can

be seen from the multitude of functions that proteins perform in nature. However, with the chosen production system, we have one additional requirement: that the product be expressed and non-toxic to the production organism.

Factors that increase toxicity and decrease the probability of good expression are hydrophobicity, positive charge and high charge density, because these are physical properties that can cause molecules to interfere with biological membranes [15, 16]. Least likely to be expressed are positively charged and amphiphilic molecules [15, 16]. Therefore, most likely to be expressed are hydrophilic molecules with not too much charge. We took these extra requirements into consideration and looked at available natural and non natural amino acid sequences to choose and design candidates for our blocks. We tried producing several different block combinations and were successful with the following combination. As a flexible block we chose the hydrophilic collagen-like C block (**Chapter 2**), which is flexible and soluble under all tested conditions, and whose sequence had been well expressed in *P.pastoris* before [9]. As a charged block, we chose the silk-like S block (**Chapter 2**) because, in its charged state, it has a moderate charge density of 1 in 8 amino acids, and it is hydrophilic and therefore likely to be expressed. But when decharged, it folds (**chapter 3**), internalizing the hydrogen bonds, and becomes hydrophobic, leading to self-assembly. In exploratory work not presented in this thesis, we also produced genes encoding combinations of blocks with higher charge density (1 in 2) and more hydrophobic amino acid residues (like leucine) but, upon induction of these genes, the *P.pastoris* cells died quickly, probably due to toxicity of the products to the production host.

Because the genes that we designed were very repetitive, we could not use the usual procedure for gene construction, namely annealing and connecting several long oligonucleotides to form a gene that then is multiplied by PCR. Because of repetitiveness, the different oligonucleotides would be very similar, and therefore annealing would occur randomly, leading to a multitude of faulty products. So, instead, we used a cloning strategy of restriction and recursive ligation (**Chapter 2**), where, after ligation, the recognition sites of the restriction endonucleases are not included in the sequence where repeated sequences or blocks are joined to each other. To this end, we created our own multiple cloning site (MCS) that enables seamless cloning and enlargement of blocks, after which any block can be coupled to any other block produced in this MCS. The DNA linking the two blocks will only code for glycine and alanine. This “toolbox” enabled us to quickly clone and couple large

blocks into genes. The ready genes were then transferred to pPIC9 and into our selected production organism *P.pastoris*, after which fermentation and purification by selective precipitation with salt and solvent led to final product yields of about a gram of product per liter of cell-free broth (**Chapter 2**). So, the first project target: “To produce novel, large, monodisperse block copolymers, with monodisperse blocks, in sufficient amounts for material testing.” succeeded, thanks to a well chosen design, efficient cloning and an efficient, well chosen expression system.

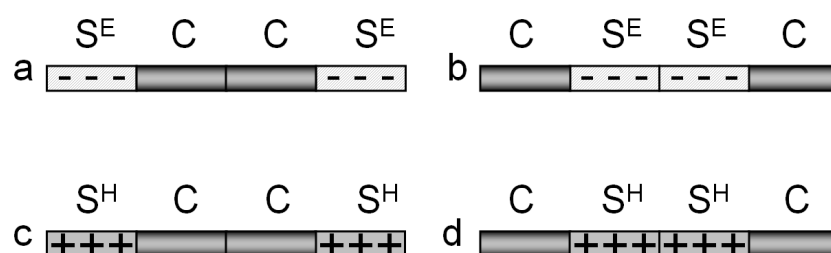


Figure 6.3 Cartoon of the produced blockcopolymers. (a) S^ECCS^E with the negatively charged, aggregating S^E blocks flanking the flexible, hydrophilic CC middle block, (b) CS^ES^EC with a negatively charged, aggregating S^ES^E middle block, flanked by hydrophilic C blocks, (c) S^HCCS^H with the positively charged, aggregating S^H blocks flanking the flexible, hydrophilic CC middle block, (d) CS^HS^HC with a positively charged, aggregating S^HS^H middle block, flanked by hydrophilic C blocks.

We produced four different triblock copolymers, respectively denoted as: CS^ES^EC, S^ECCS^E, CS^HS^HC and S^HCCS^H (**Chapter 2**) (**Fig. 6.3**). Both of the S^E containing products were negatively charged with glutamic acid residues (E) in the assembling S^E blocks. Both of the S^H containing products were positively charged with histidine residues (H) in the assembling S^H blocks. These block copolymers were supposed to self-assemble when shifting the pH appropriately, or to co-assemble into molecular networks (**Fig. 6.4 a-c**) when the oppositely charged protein polymers were mixed at moderate pH. The (GAGAGAGE)_n amino acid sequence of the S^E block had been described before [17]. Similar amino acid sequences with a different eighth amino acid residue instead of glutamic acid (E) have been described before too [18], but replacing the glutamic acid by histidine in this sequence, to create a positively charged counterpart to the S^E block was new, especially when expressed as part of an entirely biosynthetic block copolymer. A patent, concerning the positively charged molecules has now been jointly filed by Philips Healthcare and the Dutch Polymer Institute (DPI). We even produced an additional set of positively charged polymers (not described in this thesis) containing lysine (K): CS^KS^KC and S^KCCS^K. We did not expect these molecules to be produced very well, yet they were, despite their relatively strong positive charge and more

hydrophobic nature. Because of their poor (water) solubility, we focused, in this thesis, on the hydrogels of S^H and S^E containing products which were more water soluble and therefore easier to handle.

Since the cloning of the C and S block containing genes, we developed a modular cloning system (not described in this thesis) which, will facilitate the creation of many more block copolymer encoding genes. It consists of a small DNA sequence containing 2 pairs of restriction endonucleases that can be used for seamless cloning as described above. The first pair of enzymes is used for enlarging blocks seamlessly to the desired length. The second enzyme pair is used to connect any block with any other block in any desired block sequence. Any newly designed block can easily be incorporated in an existing design or combined with blocks to encode novel block copolymers. This genetic adapter enabled us to clone an additional 9 different genes encoding 9 different block copolymers and will be an asset in facile cloning of many new genes encoding new protein polymers.

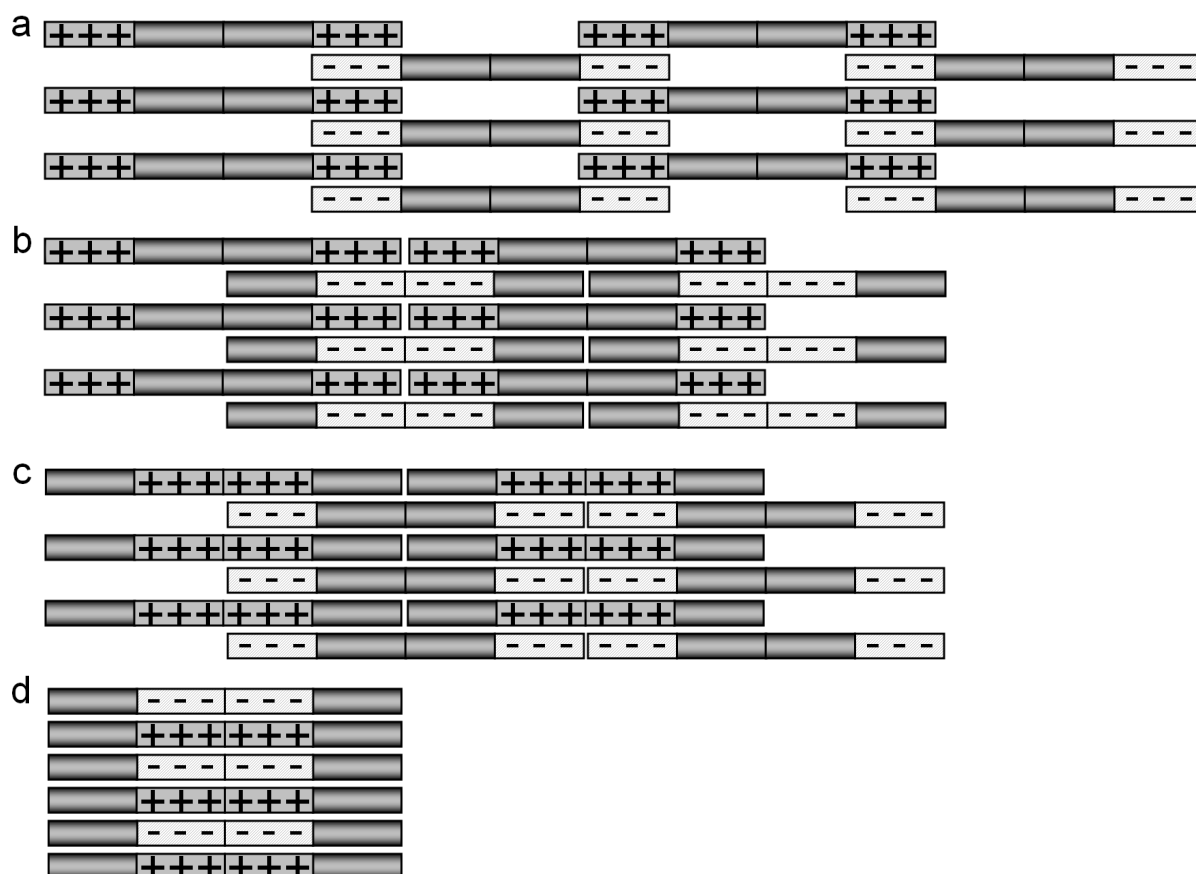


Figure 6.4 Cartoon of possible networks from combining the different oppositely charged products: (a) $S^H CCS^H$ and $S^E CCS^E$. (b) $S^H CCS^H$ and $CS^E S^E C$. (c) $CS^H S^H C$ and $S^E CCS^E$. (d) The combination of $CS^H S^H C$ with $CS^E S^E C$ can not lead to such a molecular network because both of the molecules have aggregating middle blocks. It is more likely that they will produce particles or long stacks.

To satisfy our curiosity, we did some exploratory experiments. We lowered the pH of 5 g/l $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ which both gelled, just like 5 g/l $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ and $\text{S}^{\text{H}}\text{CCS}^{\text{H}}$ gelled when the pH was raised. The four possible 1:1 mixes between positively and negatively charged products also gelled, as expected. Mixtures of equally charged products did not gel. To our satisfaction, these experiments confirmed the intended behavior that charge compensation would lead to aggregation and network formation.

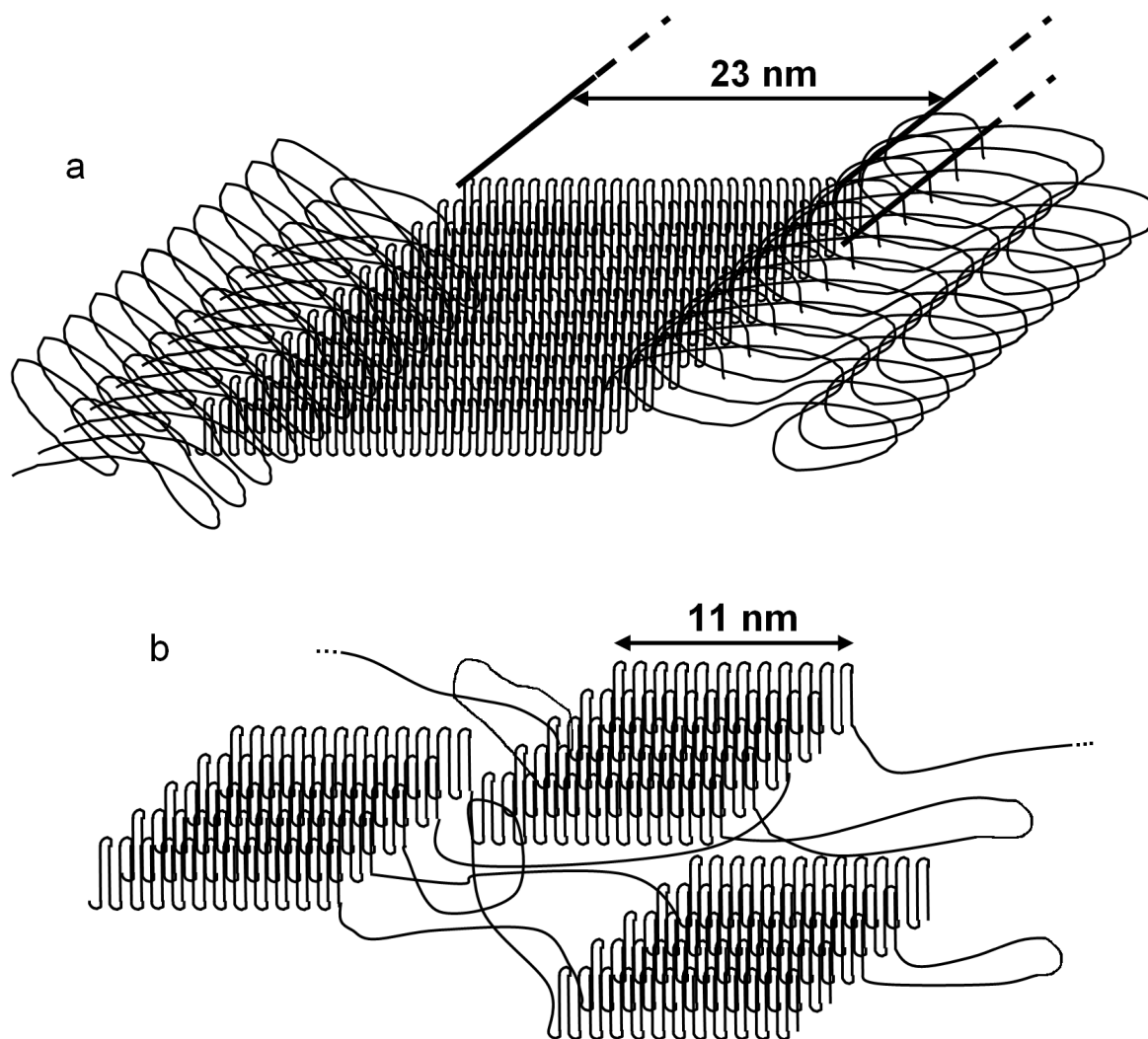


Figure 6.5 (a) Impression of expected $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ structure at low pH: a stack of β -sheets, flanked by random coils. (b) Impression of expected $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ structure at low pH: smaller crystalline β -sheet stacks, connected by the CC middle block.

We took a more systematic approach to test the idea that monodispersity of molecules and blocks would lead to neat self-assembly. First we characterized the self-assemblies of the separate products, then we studied co-assemblies of the single components combined with different oppositely charged polyelectrolytes, and finally we studied the co-assembled

mixtures of our positively and negatively charged protein block copolymers. These studies were carried out in different depth. To gain a basic understanding of the behavior of the molecules, $CS^E S^E C$ was chosen as the model molecule and its self-assembly into ribbons and gels was studied in-depth using CD spectrometry, MD modeling, SAXS, and mechanical spectroscopy. The study of the other products, or combinations of products with each other or oppositely charged polyelectrolytes was more limited for the sake of efficiency.

First of all, the separate products $CS^E S^E C$ and $S^E CCS^E$ were subjected to an extensive investigation (**Chapters 3, 4 and 5**). Let us compare our expectations of these molecules to the results. When lowering the pH, we expected them both to self-assemble, forming a structure (**Fig. 6.5**) similar to the crystalline stack of β -sheets suggested for a $(GAGAGAGE)_n$ (with $n = 10$ or 20) middle block, flanked by covalently bound PEO [19]. In a mixture of methanol and formic acid, the polypeptide part (middle block) of these molecules folded into β -sheets that stacked into ribbons that were deposited from this mixture of organic solvents.

Especially for the $CS^E S^E C$ molecule, containing also a middle block with the same octapeptide repeat ($n = 48$) we expected a very similar single stack of molecules resulting in long fibrils with a crystalline $S^E S^E$ block core and a hydrophilic C block corona (**Fig. 6.5 a**), while for the $S^E CCS^E$ molecule with self-assembling S^E outer blocks ($n = 24$) flanking a hydrophilic CC middle block, we expected to see many smaller self-assembled stacks, linked by the hydrophilic CC middle block (**Fig. 6.5 b**) comparable to a network of flower-like micelles. In contrast to the amorphous core of a flower-like micelle, the core (stack) of our structure would be crystalline. We expected these S^E outer block stacks to have only half the width of the $S^E S^E$ middle block stacks (**Fig. 6.5**), because a stack would consist of only S^E outer blocks and not double-sized $S^E S^E$ middle blocks.

At first glance, (AFM, TEM, cryo-TEM), (**Chapter 3**) we found that $CS^E S^E C$ indeed produced the nanoscopic structures that we expected: μm long fibrils of uniform width, consisting of a crystalline core and a hydrophilic corona. The crystalline core had, as expected, the shape of a ribbon. However, the width of the ribbon (SAXS **Chapter 3**) was only half of what we expected for a stack of β -sheets [19]. Also, the CD spectrum was inconsistent with β -sheets. Therefore, the structure of the fibril core formed in water could not be a stack of β -sheets like the one found for fibrils formed in organic solvent [17, 19]. With MD modeling (**Chapter 3**), performed by Peter Bolhuis and Marieke Schor (University of Amsterdam), we found only one orderly structure for the uncharged S^E blocks that was stable

in water: an unusual β -roll (**Fig. 6.6 a**). It has only four amino acid residues per β -strand and four amino acid residues per unusual turn. It is tightly wound, with the backbones of the strands against each other, and the methyl side groups of the alanine residues sticking outwards, forming “rows of knobs”. Usually, β -rolls, occurring in natural proteins across different species and functions, have space inside for residue side chains [20-25], and most of them are right-handed and have more spacious, Ca^{2+} binding turns [20, 21, 23, 25]. Some however, like our β -roll, are left-handed [24] and some do not bind Ca^{2+} [22, 24], but still, the turns that we find are different. This unusual β -roll is capable of stacking hydrophobically (**Fig. 6.6 b**) by fitting the “knobs and holes” formed by the S^{E} methyl groups, similarly to what was suggested for β -sheets with the same aminoacid sequence [17, 19]. According to MD modeling (**Chapter 3**), stacking stabilizes the β -roll structure even further. Moreover, a stack of such β -rolls (**Fig. 6.7 a**) has dimensions consistent with our SAXS measurements (**Chapter 3**). It seems reasonable to suppose that the tightly wound, unusual β -roll caused the unusual CD spectrum. We may conclude that the solvent from which the S^{E} block is crystallized has a major influence on its final structure and leads to either the formation of a β -sheet (from organic solvent) or to another structure, probably the β -roll (from water). The free energy difference between them is probably small.

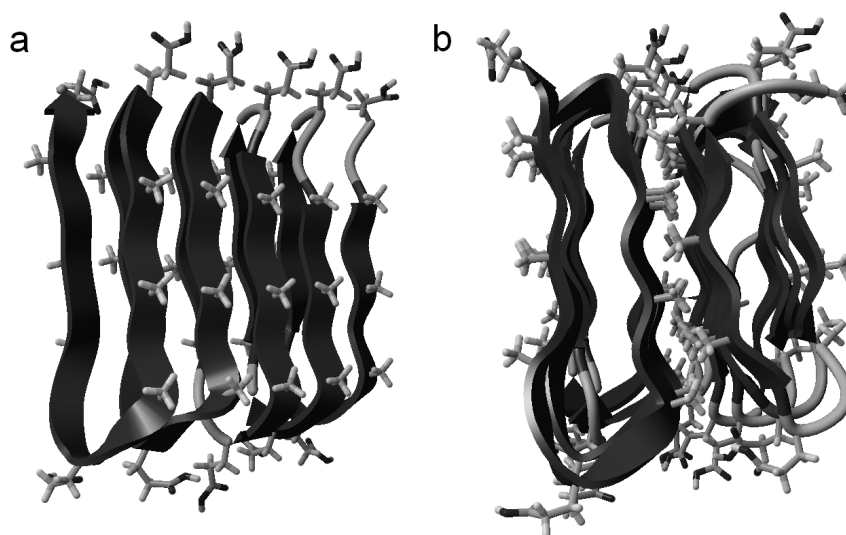


Figure 6.6 (a) Model of the β -roll conformation of an S^{E} block analogue $(\text{GAGAGAGE})_{10}$. (b) Two of such β -rolls stacked as they would be in a fibril core.

We assumed that the β -roll fibril structure contributed to the stable CD spectrum of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ at low pH. The time dependent CD signal, i.e. following the conformational

transition (**Chapter 4**) from free molecules to fibril, revealed that, after acidification, the amount of fibril in a sample increased monoexponentially in time, indicating first order kinetics for fibril formation. Comparing the monoexponential curves of several $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ concentrations revealed an initial fibril growth indicating a second order reaction. We explained this with nuclei in our samples that originated from the protein stock solution combined with a nucleation-and-growth mechanism for fibril formation (**Chapter 4**) in which nucleation is extremely slow compared to fibril growth. Seeding a sample with nuclei might be used in the future to speed up gel formation for applications that demand quick gelling.

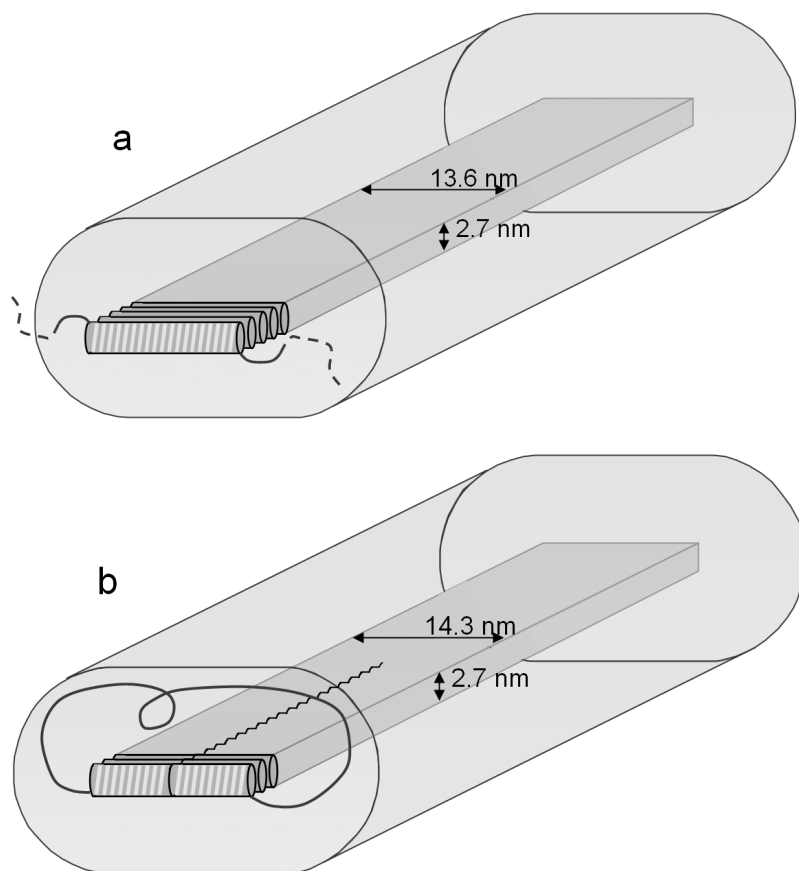


Figure 6.7 (a) Cartoon of a $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ fibril, where the middle $\text{S}^{\text{E}}\text{S}^{\text{E}}$ block forms the core of stacked β -rolls and the C blocks form the hydrophilic corona (b) Cartoon of a $\text{S}^{\text{E}}\text{C C S}^{\text{E}}$ fibril, where the separate S^{E} blocks form a double row of stacked β -rolls in the core and the CC middle block forms a corona of hydrophilic loops.

For $\text{S}^{\text{E}}\text{C C S}^{\text{E}}$ we found results very similar to those for $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$. Much to our surprise, there was hardly any difference between the structures formed (**Chapter 3**) by $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and $\text{S}^{\text{E}}\text{C C S}^{\text{E}}$. They had the same size and shape: μm long fibrils with a crystalline core with a hydrophilic corona, where we had expected smaller crystallites crosslinked by flexible CC

middle blocks. The crystalline core had almost the same size (SAXS), (**Chapter 3**) as for $CS^E S^E C$, which we explained with a tentative zipper-like structure (**Chapter 3**), (**Fig. 6.7 b**) consisting of two parallel stacks of S^E block β -rolls that have hydrogen bonds between them. For a single $S^E CCS^E$ fibril, the S^E blocks of one molecule sit in the same fibril core with one block in each stack constituting the core. The CC middle blocks then must form loops. This is actually not so different from classical block copolymers with sticky end blocks, (so-called telechelics) that form flower-like micelles [26], also consisting of a hydrophobic core and a corona of loops. The main difference (apart from size) between classical block copolymers and our molecules is that our hydrophobic core is crystalline and self-assembles directionally, resulting in fibrils, as opposed to the amorphous and isotropic hydrophobic cores of flower-like micelles [26]. At higher concentrations, classical block copolymers with sticky end blocks can form networks of flower-like micelles [26] when, for an individual molecule, it becomes possible for its sticky ends to sit in different hydrophobic cores, connecting the different micelles to form a network. We have to take into account the possibility of a similar concentration-dependent structure for the $S^E CCS^E$ molecules (**Fig. 6.8**), although after formation, this structure is kinetically trapped, in contrast to the more dynamic nature of most classical block copolymer micelles. Our molecules mainly have both of their sticky ends in the same fibril core. If a few of our fibrils are growing parallel to each other, an incoming molecule might get its sticky ends caught in two different cores, that run parallel to each other. In cryo-TEM (**Chapter 3**) we see cores running parallel to each other, but no evidence of bundling. This is because all CC middle blocks are water-swollen and equal in length, leaving the fibril cores equally spaced. However, when dried, like for TEM imaging (**Chapter 3**), the water swollen CC blocks collapse. The molecules that have their two “sticky” blocks in different cores then pull the cores together, forming bundles. This is a mechanism that can only occur for a molecule with sticky end blocks and not for a molecule with only one sticky middle block. Consequently, we see bundling for dried $S^E CCS^E$ but not for dried $CS^E S^E C$ (**Chapter 3**).

Although, at low pH, the S^E block forms a β -roll instead of a β -sheet, it seems that our expectation for the structure of $S^E CCS^E$, as depicted in **Figure 6.5**, was partially correct. We do have crystalline stacks that are to some extent crosslinked by CC middle blocks, but instead of small crystallites, we have long ribbons consisting of two adjacent large crystalline stacks in a zipper-like structure. We expect that the higher the protein concentration, the more

CC block crosslinks will occur between the stacks, and the closer the system will meet our initial expectation of a molecular network. Still, at high concentrations, large stacks, be it heavily crosslinked by CC middle blocks, seem more likely than small crystallites.

Gelling (without extra seeding) was generally slow, so we only studied concentration dependent mechanical properties for $CS^E S^E C$ gels. The gels that we produced by acidifying $CS^E S^E C$ solutions, ranging from 1 g/l to 15 g/l (**Chapter 5**), had the highest moduli ever recorded for protein hydrogels at comparable protein concentrations [27-29]. The reason why they formed such stiff gels at low concentrations is that the self-assembled fibrils are rigid filaments, which have a high aspect ratio and are crosslinked, allowing a small amount of material to form a rigid continuous structure throughout the sample volume. The gels were even stiffer than highly crosslinked actin gels [27-29]. Like actin gels, our gels displayed strain hardening and a modulus/concentration relationship indicative of rigid filaments in a crosslinked network [27]. The nature of the crosslinks between the filaments must be physical, and probably occurs between the surfaces of the ribbons that are densely covered with uncharged glutamic acid residues (carboxylic acid groups), all capable of forming hydrogen bonds with each other. Most likely, massive hydrogen bonding between the surfaces of crossing ribbons causes the physical crosslinking which is indirectly observed in our mechanical spectroscopy measurements.

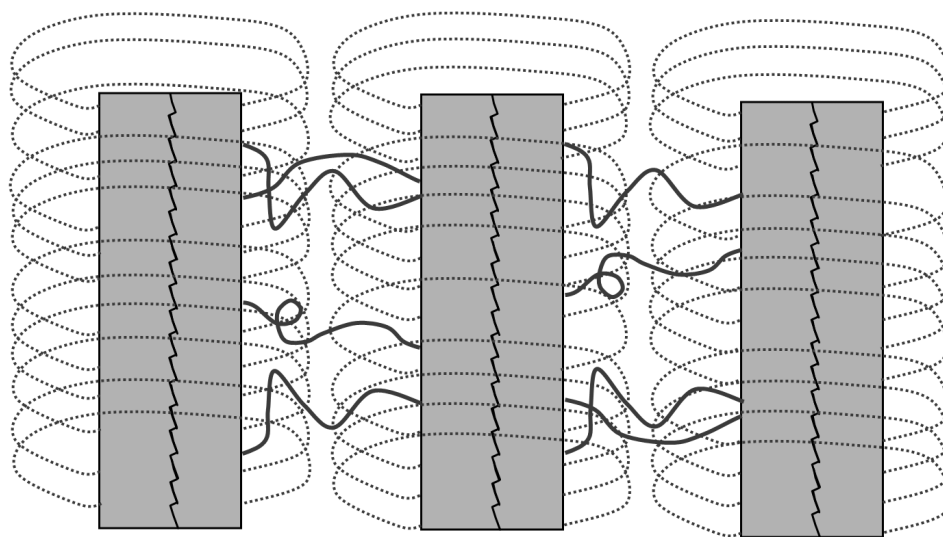


Figure 6.8 Cartoon of parallel $S^E CCS^E$ fibrils, where the separate S^E blocks form a double row of stacked β -rolls in the core and the CC middle block forms a corona of hydrophilic loops. Occasionally the S^E blocks sit in different fibril cores giving the fibrils extra flexible crosslinks.

At 8 g/l of product, $S^E CCS^E$ appeared to form an even stiffer gel than $CS^E S^E C$ (**Chapter 5**). This is probably due to extra (flexible) crosslinking between the crystalline stacks (**Fig. 6.8**), by the CC middle blocks of molecules with S^E blocks in different crystalline cores. When strained to failure, the $CS^E S^E C$ gel fractured while the $S^E CCS^E$ tore: the modulus of the latter slowly decreased with increasing deformation rather than dropping steeply as was the case with the former (**Chapter 5**). This is again consistent with the proposed zipper-like structure for $S^E CCS^E$ which might come apart according to the following tearing mechanisms. When the crystalline stacks break due to deformation, $S^E CCS^E$ still has a “softer” network in place running via the CC blocks and remaining stacks. Molecules bridging two stacks are then extended until one of their sticky blocks is slowly pulled out from a stack. This mechanism is not possible for $CS^E S^E C$, for which one molecule can only occupy one stack and therefore $CS^E S^E C$ gels are brittle, breaking all of a sudden.

Both $CS^H S^H C$ and $S^H CCS^H$, at high pH, produced much weaker gels, with about ten times lower modulus (**Chapter 2**) than the S^E containing products at the same concentration. This could be explained with the observation (microscopy **Chapter 3**) that the S^H containing products form gels of both stiff fibrils and kinetically trapped random aggregates (consistent with CD and SAXS measurements (**Chapter 3**)). The formed fibrils are very similar to the ones formed by the S^E containing products, but since a large part of the molecular population aggregates randomly, one obtains for a given amount of material, a much lower amount of stiff fibrils than in the case of S^E , while the weakly connected random aggregates prevent the stiff fibrils to approach each other to form strong crosslinks. Lack of strain hardening of these gels indeed indicated that little or no crosslinks were present.

After studying the self-assembly of the single products, induced by shifting the pH, we mixed our molecules with oppositely charged polyelectrolytes to see what kind of structures would be produced by co-assembly (**Chapter 3**). In contrast to pH induced aggregation of $CS^H S^H C$, co-assembly with negatively charged polymers did not lead to random aggregation but only to orderly fibrils that were seen with TEM, cryo-TEM and SAXS (**Chapter 3**). The difference between the two situations is that at high pH the histidine residues are intrinsically uncharged and might therefore stick randomly to each other, while when co-assembling at moderate pH, the histidine residues are still charged and have to be presented on the protein surface for them to be neutralized by the other polyelectrolyte.

Co-assembly with different charge-compensating polyelectrolytes: long PAA, short conducting POWT, and supramolecular $[\text{Zn}^{2+}\text{L}_2\text{EO}_4^{4-}]$ (**Chapter 3 section III**) and self-assembly at high pH (**Chapter 3 section II**) led to different structures. PAA and pH induced fibrils could contain β -rolls, consistent with their size determined from cryo-TEM. The corresponding CD spectra could be convoluted with those of random structures. $[\text{Zn}^{2+}\text{L}_2\text{EO}_4^{4-}]$ induced fibrils were consistent in size with a stack of β -sheets (**Fig. 6.5**) [17, 19] although the CD spectrum suggested a combination of β -sheet with α -helix and β -roll structures or another unresolved structure. It seems that these structures are energetically close to each other and might therefore be interconvertable. For a fibril, conversion of the β -rolls (**Fig. 6.7 a**) into β -sheets (**Fig. 6.5 a**) would lead to a change of shape of the fibril, becoming about twice as short and twice as wide. This might open possibilities for construction of molecular actuators. Possibly such a conformational transition might likely explain the enormous increase in modulus of the 1 g/l $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ gel as a function of temperature (**Chapter 5**), supposing that due to the raise in temperature, the fibrils became shorter, wider and stiffer. This can however not be said with any certainty until more structural characterization is done.

Finally, we mixed positively and negatively charged protein polymers (**Chapter 3**), after first testing their behavior as single products in response to pH or oppositely charged polyelectrolytes. Only 1:1 charge ratios of positively and negatively charged protein polymers were studied. They formed mainly random aggregates with some linear structures. Still this limited set of mixtures displayed a wealth of different material properties with varying moduli and elasticity (data not shown), depending on pH and total protein concentration. It is likely that more control can be obtained by choosing more sophisticated co-assembly protocols than simply mixing. E.g. one could mix at high ionic strength and then slowly dialyze out the salt. Much is still to be done, in terms of mixing and material processing, to explore the total range of material properties obtainable for our multi component protein gels. This should make it possible, for example, to produce gels made solely from biocompatible proteins that are tailored to specific biological requirements, which is ideal for, e.g., tissue engineering purposes with a high demand on material properties like engineering heart tissue.

Generally there are two methods for creating a biological tissue. The first [30-34] is by producing a scaffold (by any means) in the absence of living cells, and seeding it with living cells afterwards, but then we have to wait for the cells to migrate and grow into the tissue. The other strategy [32, 35, 36] is by mixing dissolved or liquid components, of which one contains

living cells, that together form a gel. The advantage is that cells are directly dispersed in the gel, leading to a homogeneous tissue, but we should try to avoid “bio-unfriendly” circumstances in the solution containing the cells, for example: a component might only be soluble at a pH where the cells do not thrive. Our multi protein polymer component gels would easily solve these issues. A gel can be easily prepared, because the separate protein polymer components are all soluble at moderate pH. Cells could be dispersed in one of the components, after which they are mixed and the gel sets, including the cells in an all-protein, biocompatible nanofibril matrix. In this way, pH trajectories that would be damaging for the cells (as used in some other cell matrix preparation methods) is prevented. Living Hela cells have already been proven to have high affinity to the C block [37] and no bio-compatibility problems are anticipated for the S blocks since silk-like GA repeats have been shown to be highly biocompatible [30], although biocompatibility still has to be ascertained for each individual molecule.

Combination of only one protein polymer component with an oppositely charged polyelectrolyte could also have such tissue engineering applications, for example co-assembly with POWT (**Chapter 3**). Because of its Zwitter ionic nature, the conductive polymer POWT co-assembles both with the negatively charged $CS^E S^E C$ and with the positively charged $CS^H S^H$ to form fibrils (**Chapter 3**), illustrating that these protein polymers might be used to act as template for nanowires made of charged, conductive polymers that normally do not self-assemble. Very stiff biocompatible gels of rigid and conductive nanowires have potential for neural tissue engineering applications [38-40], in which biocompatibility and conductivity are two highly desired properties. These properties are also highly desirable for coatings of neural electrodes for implantation. Gels of electrically conductive nanowires are also attractive for cardiac tissue engineering [38-40], since synchronous beating of heart cells is physiologically regulated by an electric pulse. Until now, most attempts to build such conductive biocompatible materials involve aromatic conductive polymers like polyaniline [38-40] but to build conductive materials and coatings, aromatic polymers are not essential. It has been shown that polyelectrolyte salts, can be protonic conductors [41-43]. If, in medical applications, long term exposure to aromatic compounds would appear to have adverse effects, a material composed of different charged protein polymers, possibly in combination with a (non aromatic) polyelectrolyte like PAA, might be a better choice.

Additionally to medical applications, molecules displaying stimulus responsiveness, protonic conductivity, and hierarchical self-assembly can lead to various functional materials [5, 6, 44] with a host of possible applications; as examples we may mention membranes for hydrogen fuel cells, photonic crystals for thin film optics and optical computers, coatings and nanowires for (bio)electronic devices, and pH responsive materials in microfluidic systems. We still have to explore the full range of hierarchically self-assembled (nano)structures that our molecules may display. Many different self-assembled structures could be achieved with our products, either as separate components, or when combined with different oppositely charged compounds [5, 6] like small amphiphiles, polyelectrolytes or supramolecular polymers. Also different physical stimuli like temperature and pH are likely to influence the architecture of the self-assemblies [5] from which even more applications might follow. Crosslinking the C blocks by their lysine residues, in any of the above structures, will lead to permanent, pH swellable gels with possibly improved (gel)material properties.

The success with which fibers can be spun from a gel of self-assembling polymers will heavily depend on the nanostructure of the gel. The single component gels have a structure analogous to **Figure 6.4 d**. Gel spinning such ribbons cannot lead to strong fibers, because aligning the ribbons into a fiber will largely lead to the strong covalent bonds running perpendicularly to the fiber, while the weaker physical interactions running along the fiber bear the load.

For spinning a gel into strong fibers, the starting material should be a molecular network as in **Figure 6.4 a**, for a maximal number of load bearing chains. Gel spinning an architecture as in **Figure 6.4 b** or **c**, in which the flexible blocks are crosslinked by their lysine residues, could also lead to strong fibers.

However, at low concentrations our molecules have the propensity to stack into long self-assembled ribbons, instead of forming molecular networks with physical crosslinks. Possibly using high concentrations of S^ECCS^E and/or S^HCCS^H will lead to more of a molecular network like in **Figure 6.8**, but still, in this structure, many molecules have both “sticky” blocks in the same ribbon core. After gel spinning, this will make the associated CC middle blocks of these molecules not bear any load.

The best candidate gel for spinning is the one with the least ribbon content and the most random structures. Because of this it is also the weakest gel, in contrast to the gels with

extremely high modulus due to their self-assembled stiff ribbons. Also it should contain the largest number of connecting flexible blocks that may bear load. Therefore the best candidate is the combination of $S^E CCS^E$ and $S^H CCS^H$. Another avenue that might be taken to create gels suitable for gel spinning is to develop a less flexible shorter middle block that cannot completely fold, preventing sticky enblocks from residing in the same core. Still, super strong fibers may be obtainable with the molecules that we produced, but for this, much more research into material processing would be needed.

For now, these molecules are most promising and suitable for creating various biocompatible gels that can reach extremely high moduli at very low polymer concentrations.

6.2 Conclusion

Important steps have been taken with respect to the production of monodisperse block copolymers with sequence containing blocks in a biotechnological fashion. We developed a genetic toolbox for quickly building genes (polymer templates) in which we can produce sequential, repetitive blocks and connect them using a modular approach. We succeeded in producing designed monodisperse triblock copolymers in amounts that were sufficient for material testing. The produced molecules were stimulus responsive to pH and oppositely charged polyelectrolytes according to our design. Depending on the pH, and on which of our produced molecules were used, either as single components or in combination with each other or oppositely charged polyelectrolytes, they formed gels consisting of either random aggregates or various self-assembled fibrils with a core/corona structure. The core/corona structures were not unlike those formed by classical block copolymers, the main difference being that we have long, self-assembled, semi-crystalline cores while block copolymer micelles usually have globular cores of amorphous polymers. Depending on how they were prepared, our gels had different material properties (strength, elasticity, possible conductivity) which could lead to applications in cardiac and neural tissue engineering. Gel formation could occur within minutes, but development of the final material properties typically took one hour to several days. For research and application purposes gel formation may be speeded up by seeding the solutions with nuclei, from which the fibrils can start to grow. Our molecules showed that monodispersity and conformational preference of amino acids favored self-assembly, although the molecules self-assembled slightly differently from the way we had expected. Still, super strong fibers may be obtainable, in which at least one molecule with the

outer “sticky” blocks: S^ECCS^E or S^HCCS^H is essential, but to do so, more research is needed into material processing.

References

- 1 Govaert, L. E., Bastiaansen, C. W. M. and Leblans, P. J. R. (1993) Stress-Strain Analysis of Oriented Polyethylene. *Polymer* **34**, 534-540
- 2 Shim, V. P. W., Lim, C. T. and Foo, K. J. (2001) Dynamic mechanical properties of fabric armour. *International Journal of Impact Engineering* **25**, 1-15
- 3 Smith, P. and Lemstra, P. J. (1980) Ultra-High-Strength Polyethylene Filaments by Solution Spinning-Drawing. *Journal of Materials Science* **15**, 505-514
- 4 Wang, A. (2001) A unified theory of wear for ultra-high molecular weight polyethylene in multi-directional sliding. *Wear* **248**, 38-47
- 5 Ikkala, O. and ten Brinke, G. (2004) Hierarchical self-assembly in polymeric complexes: Towards functional materials. *Chemical Communications*, 2131-2137
- 6 Faul, C. F. J. and Antonietti, M. (2003) Ionic self-assembly: Facile synthesis of supramolecular materials. *Advanced Materials* **15**, 673-683
- 7 Heslot, H. (1998) Artificial fibrous proteins: A review. *Biochimie* **80**, 19-31
- 8 Cappello, J., Crissman, J., Dorman, M., Mikolajczak, M., Textor, G., Marquet, M. and Ferrari, F. (1990) Genetic-Engineering Of Structural Protein Polymers. *Biotechnology Progress* **6**, 198-202
- 9 Werten, M. W. T., Wisselink, W. H., van den Bosch, T. J. J., de Bruin, E. C. and de Wolf, F. A. (2001) Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Engineering* **14**, 447-454
- 10 Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) *Pichia-Pastoris* as a Host System for Transformations. *Molecular and Cellular Biology* **5**, 3376-3385
- 11 Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., Sreekrishna, K. and Henwood, C. A. (1991) Production of Mouse Epidermal Growth-Factor in Yeast - High-Level Secretion Using *Pichia-Pastoris* Strains Containing Multiple Gene Copies. *Gene* **105**, 205-212
- 12 Cregg, J. M., Cereghino, J. L., Shi, J. Y. and Higgins, D. R. (2000) Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology* **16**, 23-52
- 13 Fahnstock, S. R. and Bedzyk, L. A. (1997) Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Applied Microbiology and Biotechnology* **47**, 33-39
- 14 Werten, M. W. T., Van den Bosch, T. J., Wind, R. D., Mooibroek, H. and De Wolf, F. A. (1999) High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**, 1087-1096
- 15 Epand, R. M. and Vogel, H. J. (1999) Diversity of antimicrobial peptides and their mechanisms of action. *Biochimica Et Biophysica Acta-Biomembranes* **1462**, 11-28
- 16 Shai, Y. (1999) Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochimica Et Biophysica Acta-Biomembranes* **1462**, 55-70
- 17 Krejchi, M. T., Atkins, E. D. T., Waddon, A. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1994) Chemical Sequence Control Of Beta-Sheet Assembly In Macromolecular Crystals Of Periodic Polypeptides. *Science* **265**, 1427-1432
- 18 Cantor, E. J., Atkins, E. D. T., Cooper, S. J., Fournier, M. J., Mason, T. L. and Tirrell, D. A. (1997) Effects of amino acid side-chain volume on chain packing in genetically engineered periodic polypeptides. *Journal Of Biochemistry* **122**, 217-225
- 19 Smeenk, J. M., Otten, M. B. J., Thies, J., Tirrell, D. A., Stunnenberg, H. G. and van Hest, J. C. M. (2005) Controlled assembly of macromolecular beta-sheet fibrils. *Angewandte Chemie-International Edition* **44**, 1968-1971

- 20 Aachmann, F. L., Svanem, B. I. G., Guntert, P., Petersen, S. B., Valla, S. and Wimmer, R. (2006) NMR structure of the R-module - A parallel beta-roll subunit from an *Azotobacter vinelandii* mannuronan C-5 epimerase. *Journal of Biological Chemistry* **281**, 7350-7356
- 21 Baumann, U. (1994) Crystal-Structure of the 50 Kda Metalloprotease from *Serratia-Marcescens*. *Journal of Molecular Biology* **242**, 244-251
- 22 Kuiper, M. J., Davies, P. L. and Walker, V. K. (2001) A theoretical model of a plant antifreeze protein from *Lolium perenne*. *Biophysical Journal* **81**, 3560-3565
- 23 Lilie, H., Haehnel, W., Rudolph, R. and Baumann, U. (2000) Folding of a synthetic parallel beta-roll protein. *Febs Letters* **470**, 173-177
- 24 Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M. and Goldman, A. (2004) The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *Embo Journal* **23**, 701-711
- 25 Yoder, M. D. and Jornak, F. (1995) Protein Motifs .3. the Parallel Beta-Helix and Other Coiled Folds. *Faseb Journal* **9**, 335-342
- 26 Xu, B., Yekta, A., Li, L., Masoumi, Z. and Winnik, M. A. (1996) The functionality of associative polymer networks: The association behavior of hydrophobically modified urethane-ethoxylate (HEUR) associative polymers in aqueous solution. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **112**, 239-250
- 27 Gardel, M. L., Shin, J. H., MacKintosh, F. C., Mahadevan, L., Matsudaira, P. and Weitz, D. A. (2004) Elastic Behavior of cross-linked and bundled actin networks. *Science* **304**, 1301-1305
- 28 Janmey, P. A., Hvidt, S., Lamb, J. and Stossel, T. P. (1990) Resemblance of Actin-Binding Protein Actin Gels to Covalently Cross-Linked Networks. *Nature* **345**, 89-92
- 29 Storm, C., Pastore, J. J., MacKintosh, F. C., Lubensky, T. C. and Janmey, P. A. (2005) Nonlinear elasticity in biological gels. *Nature* **435**, 191-194
- 30 Altman, G. H., Diaz, F., Jakuba, C., Calabro, T., Horan, R. L., Chen, J. S., Lu, H., Richmond, J. and Kaplan, D. L. (2003) Silk-based biomaterials. *Biomaterials* **24**, 401-416
- 31 Matthews, J. A., Wnek, G. E., Simpson, D. G. and Bowlin, G. L. (2002) Electrospinning of collagen nanofibers. *Biomacromolecules* **3**, 232-238
- 32 Hutmacher, D. W. (2000) Scaffolds in tissue engineering bone and cartilage. *Biomaterials* **21**, 2529-2543
- 33 Madihally, S. V. and Matthew, H. W. T. (1999) Porous chitosan scaffolds for tissue engineering. *Biomaterials* **20**, 1133-1142
- 34 Freed, L. E., Vunjaknovakovic, G., Biron, R. J., Eagles, D. B., Lesnoy, D. C., Barlow, S. K. and Langer, R. (1994) Biodegradable Polymer Scaffolds for Tissue Engineering. *Bio-Technology* **12**, 689-693
- 35 Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A. and Stupp, S. I. (2004) Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science* **303**, 1352-1355
- 36 Drury, J. L. and Mooney, D. J. (2003) Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials* **24**, 4337-4351
- 37 Rozkiewicz, D. I., Kraan, Y., Werten, M. W. T., de Wolf, F. A., Subramaniam, V., Ravoo, B. J. and Reinhoudt, D. N. (2006) Covalent microcontact printing of proteins for cell patterning. *Chemistry-a European Journal* **12**, 6290-6297
- 38 Guo, Y., Li, M. Y., Mylonakis, A., Han, J. J., MacDiarmid, A. G., Chen, X. S., Lelkes, P. I. and Wei, Y. (2007) Electroactive oligoaniline-containing self-assembled monolayers for tissue engineering applications. *Biomacromolecules* **8**, 3025-3034
- 39 Huang, L. H., Hu, J., Lang, L., Wang, X., Zhang, P. B., Jing, X. B., Wang, X. H., Chen, X. S., Lelkes, P. I., MacDiarmid, A. G. and Wei, Y. (2007) Synthesis and characterization of electroactive and biodegradable ABA block copolymer of polylactide and aniline pentamer. *Biomaterials* **28**, 1741-1751
- 40 Li, M. Y., Guo, Y., Wei, Y., MacDiarmid, A. G. and Lelkes, P. I. (2006) Electrospinning polyaniline-contained gelatin nanofibers for tissue engineering applications. *Biomaterials* **27**, 2705-2715

- 41 Rikukawa, M. and Sanui, K. (2000) Proton-conducting polymer electrolyte membranes based on hydrocarbon polymers. *Progress in Polymer Science* **25**, 1463-1502
- 42 Antonietti, M., Maskos, M., Kremer, F. and Blum, G. (1996) Dielectric relaxation and conductivity in polyelectrolyte-surfactant complexes. *Acta Polymerica* **47**, 460-465
- 43 Antonietti, M., Neese, M., Blum, G. and Kremer, F. (1996) Dielectric and mechanic relaxation in polyelectrolyte-supported bilayer stacks: A model for the dynamics of membranes? *Langmuir* **12**, 4436-4441
- 44 Ikkala, O. and ten Brinke, G. (2002) Functional materials based on self-assembly of polymeric supramolecules. *Science* **295**, 2407-2409

Summary

The research described in this thesis concerns the design, biotechnological production, and physiochemical study of large water-soluble (monodisperse) protein triblock-copolymers with sequential blocks, some of which are positively or negatively charged and self-assemble in response to a change in pH or co-assemble in response to oppositely charged polyelectrolytes (including each other). Such molecules displaying controlled self-assembly may lead to new biocompatible nano-structured materials like nano-wires, gels and fibers, with unusual material properties and potential technical and biomedical applications. First, the production and purification as enabling technology is described but the focus of this thesis is on the physiochemical behavior of the produced molecules, including the nano-structures formed, the kinetics of the structure formation and the material properties of the macroscopic gels resulting from the nano-structures.

For producing monodisperse block copolymers with sequential blocks, we chose the natural protein production machinery of living cells, because monodispersity and sequentiality are hallmarks of proteins. First, DNA encoding various polypeptide blocks of the block copolymers was designed. Then, using a modular cloning approach, DNA blocks were built, enlarged to desired block size and connected to form whole genes that encode polypeptide block copolymers. The genes were transferred to the production host, the yeast *Pichia pastoris*, which when induced, secreted the various protein block copolymers production yields in the gram per liter range, such that various applications of these promising biomaterials become possible. However, choosing this biotechnological approach limited the polymer design to the 20 natural amino acids, but seen the large variation in structure and function of natural proteins, this, in practice, forms hardly any limitation.

Three different nature-inspired poly-peptide blocks were used in the block copolymer designs: two very similar, but oppositely charged silk-like blocks, and one largely uncharged collagen like block. The silk-like blocks (“S”) consist of an octapeptide glycine (G) alanine(A) repeat (GAGAGAGX)₂₄, in which the X position is occupied either by a positively charged histidine residue (H) or a negatively charged glutamic acid residue (E), resulting in an either positively or negatively charged silk-like block: “S^H” or “S^E” respectively, the theoretical pKa being 7 and 4.3 respectively. These blocks are supposed to self- or co-assemble upon charge neutralization or compensation. The collagen-like blocks (“C”)

contains mainly glycine, proline, polar amino acids and a small number of charged residues. It is supposed to form a hydrophilic random coil in most aqueous environments. The four final products, all 802 amino acids long, were either positively charged, 66.1 kDa molecules denoted as S^HCCS^H and CS^HS^HC or negatively charged 65.7 kDa molecules denoted as S^ECCS^E and CS^ES^EC . The four different molecules showed different behavior depending on charge and block order. Three aspects of these molecules were studied: nano-structures, kinetics of structure formation, and material properties of pH change induced gels.

Nano-structures, were investigated with a broad range of techniques including: transmission electron microscopy (TEM) cryogenic (cryo-)TEM, atomic force microscopy (AFM), CD spectrometry, small angle X-ray scattering (SAXS) and molecular dynamics modeling (MD). The four different polymers were triggered to form structures and 12 of these self- or co-assembled nano-structures were examined: 1-4) the four pH- induced self-assemblies, 5-7) CS^HS^HC with poly acrylic acid (PAA), or with a metal bis-ligand supramolecular polymer denoted as $Zn-L_2(EO)_4$, or with the (conducting) polythiophene (POWT) that was chemically modified to be zwitterionic, 8) CS^ES^EC with POWT, 9-12) the four possible mixtures of the four different protein block copolymers. Except for the mixtures of protein polymers, which seemed to form kinetically trapped molecular networks, the self- and co-assemblies formed well defined μm long nanoribbons with a hydrophilic C block corona. The core structure, depended on the protein block copolymer used, and on the mode of charge compensation like pH and/or the type of polyelectrolyte used. Interesting features are: the unusual β -roll, predicted with MD modeling for the S^E block in the self-assembled ribbon core at low pH, the CC middle blocks of S^ECCS^E and S^HCCS^H forming loops, analogous to flower-like micelles, and the templating of the conductive polymer POWT onto nanoribbons, that might have applications as nanowires.

Kinetics of structure formation was only followed for CS^ES^EC at low pH. CD spectroscopy at 200 nm was used to follow the conformational change associated with ribbon formation of CS^ES^EC in time, in dilute, acidified solutions, revealing a nucleation and growth mechanism for CS^ES^EC ribbons under a critical pH of approximately 4.5. Dynamic light scattering (DLS) measurements revealed that when the pH was increased, the ribbons dissolved and the formed gels melted, but they only did so above pH 5.4 which is much higher than the critical pH of ribbon formation. This pH region in which ribbons do not form,

nor dissolve suggests a kinetic barrier to ribbon formation. The purified and freeze-dried $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ appeared to contain nuclei from which ribbon growth could start.

To study the effect of block charge and block order, on the mechanical properties of self-assembling block copolymer hydrogels, we tested the physical behavior of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$, $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$, and $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$. Dynamic mechanical spectroscopy revealed differences in gelling kinetics and mechanical properties of the three different polymers. Remarkably, the S^{E} containing polymer gels displayed non linear elasticity comparable to that of the actin gels and other biological gels. Consequently we see that like actin gels our gels consist of semi-flexible fibrils (nano-tapes). Moreover, exceptionally high dynamic elasticity moduli, exceeding 40 kPa, were reached already at concentrations as low as 1.5 wt%, without any additional crosslinking agent. Such highly rigid yet dilute gels are rare and sought after. $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ gels were relatively weak and formed slowly. Additionally we studied the effect of temperature on the mechanical properties of a $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ gel.

The molecules self- and co-assembled into various nano-structures constituting various transparent gels, some of which extremely rigid. These structures and gels have potential biomedical and technological applications. The biotechnological approach for producing these four different monodisperse, sequential block copolymers, yielded amounts that make applications possible and presents an opportunity for the design and production of many more monodisperse and sequential (protein) block copolymers for building other nano-structures.

Samenvatting

Het onderzoek beschreven in dit proefschrift betreft het ontwerp, de biotechnologische productie en de fysisch-chemische studie, van grote, wateroplosbare, monodisperse triblokcopolymeren met ongeladen hydrofiele blokken en positief of negatief geladen blokken die kunnen zelfassembleren in reactie op een verandering in de pH of coassembleren met tegenovergesteld geladen polyelectrolyten (elkaar meegerekend). Zulk soort moleculen, die gecontroleerd zelfassembleren, zouden kunnen leiden tot nieuwe, zelfassemblede biocompatibele materialen zoals nanodraadjes, gelen en vezels met ongewone materiaaleigenschappen en potentieele technische en biomedische toepassingen.

In eerste instantie wordt de productie en zuivering van deze moleculen beschreven, maar de nadruk van dit proefschrift ligt vooral op het fysisch-chemische gedrag van de geproduceerde moleculen. Dit houdt in: de nanostructuren die ze vormen, de bijbehorende vormingskinetiek en de materiaaleigenschappen van de macroscopische gelen bestaande uit deze nanostructuren.

Om monodisperse blokcopolymeren met sequentiële blokken te produceren hebben we de natuurlijke eiwit productie machinerie van levende cellen verkozen omdat monodispersiteit en een sequentialiteit typische kenmerken zijn van biologisch geproduceerde eiwitten die met bij voorbeeld organische synthese zeer moeilijk te evenaren zijn. De ontwikkeling ging als volgt in zijn werk. Eerst werden korte DNA sequenties, coderend voor verschillende aminozuursequenties, ontworpen. Vervolgens werden, gebruikmakend van een modulaire cloneringsstrategie, deze DNA sequenties verlengd tot een gewenste bloklengte. Deze DNA blokken werden daarna op een vergelijkbare wijze geschakeld tot genen die coderen voor polypeptide blokcopolymeren. De genen werden in een productieorganisme, de gist *Pichia pastoris*, geplaatst, die tijdens inductie, meer dan een gram per liter van de diverse producten bleek uit te scheiden. De uitgescheiden hoeveelheden zijn genoeg om diverse hoogwaardige toepassingen mogelijk te maken zoals biomaterialen of nanodeeltjes voor medische toepassingen. Het ontwerp van eiwitblokcopolymeren is in bouwstenen grotendeels begrensd tot de 20 natuurlijke aminozuren. Dit hoeft echter geen grote beperking te zijn in de mogelijke producten, gezien de enorme diversiteit in structuur en functie van natuurlijke eiwitten.

In de ontworpen triblokcopolymeren werden drie verschillende blokken gebruikt. Twee waren er sterk gelijkend, maar tegenovergesteld geladen. Ze hadden een zijdeachtige

aminozuursequentie (GAGAGAGX)₂₄, waarin G glycine is, A alanine en X het geladen aminozuurresidu: ofwel positief geladen histidine (H) ofwel negatief geladen glutaminezuur (E). De blokken werden respectievelijk “S^H” en “S^E” genoemd met een respectievelijke theoretische pK_a van 7 en 4.3. Deze blokken zijn ontworpen om te zelfassembleren in reactie op pH en/of tegenovergestelde lading. Het derde gebruikte blok (“C”) heeft een collageenachtige sequentie en bezit vooral polaire grepen en prolineresiduen. In tegenstelling tot veel andere collageenachtige sequenties, vormt deze aminozuursequentie in water geen gel, maar blijft hij in de vorm van een willekeurige kluwen goed opgelost bij diverse temperaturen en pH.

De vier uiteindelijke triblokcopolymeren, allemaal 802 aminozuren lang, waren de positief geladen S^HCCS^H en CS^HS^HC en de negatief geladen S^ECCS^E en CS^ES^EC. Ze vertoonden verschillend gedrag, afhankelijk van lading en van blokvolgorde. Drie aspecten van deze moleculen werden bestudeerd: nanostructuren, kinetiek van structuurvorming en materiaaleigenschappen van hun pH geïnduceerde gelen.

De nanostructuren werden bestudeerd met behulp van vele technieken waaronder: transmissie elektronen microscopie (TEM), ook van bevroren monsters (cryo-TEM), atomaire kracht microscopie (AFM), circulair dichroïsme (CD) spectrometrie, kleine hoek röntgen verstrooiing (SAXS) en moleculaire dynamica modelering (MD). Interessante vondsten waren onder andere de bijzondere β -rol die met MD voorspeld werd voor het S^E blok dat bij lage pH zelfassembleert en stapelt tot lange linten en het middelste CC blok van S^ECCS^E lussen vormt op de gezelfassembleerde linten, analoog aan de lussen in bloemvormige micellen. Verder kon, behalve met een pH verschuiving, lintvorming ook geïnduceerd worden door de toevoeging van een tegenovergesteld geladen polyelectrolyt zoals het Zwitterionische POWT, dat coassembleerde met de eiwitblokcopolymeren tot linten met een dun laagje POWT erop. Deze zouden vanwege de eigenschappen van POWT mogelijk toepassing kunnen vinden als elektrisch geleidende nanodraden.

De kinetiek van de structuurformatie is alleen bestudeerd voor het CS^ES^EC eiwit bij lage pH. Hiervoor werd het CD signaal bij 200 nm gebruikt om de conformatieverandering van het molecuul in waterige oplossing en bij lage pH van ongeordende toestand naar lint te volgen. De linten bleken zich te vormen volgens een nucleatie en groei kinetiek onder een kritische pH van 4.5. Dynamische lichtverstrooiing (DLS) liet zien dat een gel van zulke linten pas volledig oplost boven een pH van 5.4. Dit is hoger dan de kritische pH van lintvorming en

suggereert dus een kinetische barrière voor lintvorming. Het gevriesdroogde materiaal waarvan uit werd gegaan bleek kiemen te bevatten van waaruit lintvorming kon beginnen.

Om het effect van bloklading en blokvolgorde op de mechanische eigenschappen van de zelfassemblede gelen te bestuderen, hebben we het fysische gedrag van $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$, $\text{S}^{\text{E}}\text{CCS}^{\text{E}}$ en $\text{CS}^{\text{H}}\text{S}^{\text{H}}\text{C}$ vergeleken. Dynamische mechanische spectroscopie liet verschillen in geleringskinetiek en mechanische eigenschappen van gelen van de drie polymeren zien. Opvallend was dat de S^{E} bevattende moleculen niet lineaire elasticiteit vertoonden dat vooral vergelijkbaar was met actine maar ook met andere gelen van biologische oorsprong. Wij zien daarom ook dat onze gelen, net als actine, uit een netwerk van semiflexibele nanofibers bestaat (in het geval van $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ dus nanolinten). Uitzonderlijk hoge elastische moduli werden door de gelen bereikt bij voorbeeld meer dan 40 kPa bij een concentratie van maar 1.5 gewichts% zonder enige additionele crosslinker.

Onder diverse omstandigheden zelfassembleden en coassembleden de moleculen met andere moleculen en met elkaar tot diverse nanostructuren die verder leidden tot macroscopische gelen, waarvan sommige extreem rigide. De gevonden structuren en gelen hebben diverse potentiële biomedische en technische toepassingen. De biotechnologische aanpak die genomen werd om deze vier monodisperse sequentiële blokcopolymeren te maken leidde tot genoeg product om toepassingen mogelijk te maken en vormt een kans om nog vele andere monodisperse sequentiële (eiwit)blokcopolymeren te ontwerpen en te produceren voor het bouwen van allerlei andere nanostructuren.

Dankwoord

Zo, en dan is het geschreven en gedaan en denk ik: “En nu nog een dankwoord maar, dat is iets heel anders dan wat ik tot nu toe geschreven heb. Er zijn een hoop mensen die ik wil bedanken ,maar hoe moet ik het beginnen?” Ik blader zo eens door een paar andere proefschriften om te kijken wat men zoal heeft geschreven, maar dat is het ook niet, althans, het is prima voor die personen die dat geschreven hebben, maar het is niet wat ik wil zeggen. Ik denk zo aan allerlei mensen die ik gedurende mijn promotie ontmoet heb, en wordt dan overspoeld door het warme gevoel dat ik eigenlijk alle vrienden en mensen die mij hebben geholpen en bijgestaan, in een grote kamer bij elkaar zou willen hebben, dat ik heel hard kon roepen “DANKJEWEL”, en dat ik armen had, zo groot dat ik iedereen tegelijk kon omhelzen. Dat is het gevoel dat ik uit wil drukken! Bij die menigte horen ook mensen die ik voor mijn promotie heb ontmoet en iedereen die door intensieve samenwerking of begeleiding mij een hoop hebben bijgebracht maar ook zeker mensen die gewoon toevallig op een cruciaal moment mij een duwtje in een bepaalde richting hebben geven. Uiteindelijk moet je het promotie zelf doen en zelf een volledig proefschrift bouwen waar je tevreden op terug kan kijken, maar zonder alle helpende handen die ik op mijn weg daar naartoe heb ontmoet, zou het nooit hebben kunnen lukken. Bedankt aan iedereen die heft bijgedragen aan waar ik nu ben. Een aantal mensen wil ik hier in het bijzonder noemen.

Ten eerste mijn co-promotor en dagelijks begeleider op A&F, Frits de Wolf, bedankt voor de vele discussies die we hebben gehad en de vrijheid die je me gaf om binnen de kaders van “eiwitblokcopolymeer” mijn gang te gaan. Je luisterde serieus naar mijn wilde ideeën en ik kon altijd rekenen op een kritische noot, een suggestie, of een “Goed gedaan.”

Mijn promotor Gerrit Eggink. In het begin was je door omstandigheden afwezig. Het meest verantwoorde besluit was toen om mij bij de vakgroep Fysische chemie en kolloïdkunde onder te brengen. Dat pakte blijkbaar goed uit. Vanaf het moment dat je terug was kon ik altijd op je rekenen. Ik heb je als een grote steun ervaren, zeker tegen het einde van mijn promotie toen spijkers met koppen geslagen moesten worden.

Mijn promotor Martien Cohen Stuart, bedankt voor het mij leren schrijven. Ik had een of andere hobbel waar ik overheen moest. Het is mij nog steeds niet helemaal duidelijk waar het

probleem lag, maar sinds dat jij er aandacht aan besteed hebt vlot het schrijven. Ik heb genoten van je enthousiaste, positieve en soms speels-artistische instelling en de leuke tijden die we met de andere AIO's hadden zowel binnen ons gebouw als op excursies barbecues en uitjes.

Aan alle drie mijn begeleiders gezamenlijk: bedankt voor het vertrouwen dat jullie in mij stelden bij mijn aanstelling voor dit ambitieuze project en voor het rekening houden met mijn persoonlijke leven. Er waren wat hobbeltjes, maar ik heb, naar mijn gevoel, met jullie een fijne tijd gehad en harmonieus kunnen samenwerken.

De mensen van het laboratorium van fysische chemie en kolloïdkunde, de AIO's, gastmedewerkers en de vaste staf, Bedankt voor een fijne tweede helft van mijn promotie met elke week wetenschappelijke discussies, en af en toe een filmavond, borrel, uitje, barbecue, of andere activiteiten. Met name Wiebe de Vos, mijn kamergenoot en paranimf, bedankt voor het mij aanhoren tijdens mijn goede en slechte momenten. Het is een wonder dat je oren niet verschrompeld zijn en eraf gevallen. Yun Yan en Feng Li, bedankt voor de avontuurlijke experimenten, het "Wat als we jou en mijn spul bij elkaar gooien?" maar dan op wetenschappelijk niveau en de wederzijds inspirerende gesprekken. Ik wil de hele groep vertellen dat ik van de geweldige tijd heb genoten.

Peter Bolhuis en Marieke Schor, bedankt voor de prettige samenwerking en de inzichten die het modeleren ons brachten. Ik denk dat we, dank zij jullie, een veel beter beeld hebben van hoe onze nanostructuren zich vormen dan we zonder jullie gehad hadden. Ik waardeer jullie zelfkritiek op het modeleren en jullie behoefte om constant het model aan de praktijk te toetsen. Ik vind dat experimentalisten veel hebben aan modelleerders als jullie.

Giuseppe Portale, bedankt voor de prettige dagen in Grenoble. Toen ik er was, was de SAXS defect, maar een goede werkbespreking maakte het mogelijk dat je de experimenten en berekeningen met succes later zelf uitvoerde. We hielden veel contact per e-mail en telefoon, en hebben zo de klus geklaard.

De bioconversie groep van A&F wilde ik graag bedanken voor een fijne eerste helft van mijn promotie. Truus, mijn kamergenoot, bedankt voor de morele steun en het aanhoren van mijn

dagelijkse problemen. Marc Werten, mijn paranimf, bedankt voor het functioneren als vraagbaak en de discussies over muziek. Elk technisch probleem scheen jij op te kunnen lossen. Je was een ware steun. Antoine Moers bedankt voor de hulpen raad bij het fermenteren en zuiveren van eiwit. Ik dank jullie en de rest van de groep en ook de andere AIO's binnen bioconversie voor de gezelligheid en het gevoel dat je deel van uitmaakt van een diverse maar toch saamhorige groep.

Ik wil graag de mensen van DPI bedanken die dit project mogelijk maakten. Piet Lemstra die het project goedkeurde, Bedankt voor je interesse en vertrouwen in mij. Sommige zaken liepen misschien niet helemaal zoals we hadden gewild, maar jou vertrouwen in mijn kunnen gaf mij een extra stuk zelfvertrouwen. Richard van den Hof, voor mij was je een tijd lang het gezicht van DPI. De ontmoetingen en correspondentie was altijd prettig. In mijn interacties met DPI als geheel had ik altijd het gevoel dat je aan mijn kant stond en je hebt me altijd geholpen als ik weer eens iets niet kon uploaden of vinden op het DPI intranet.

Van de mensen die bijgedragen hebben aan mijn vorming voor mijn promotie wil ik vooral Ed van Niel en , Ana Lopez Contreras bedanken. Jullie hebben mij tijdens mijn studententijd verschillende vaardigheden bijgebracht die cruciaal bleken tijdens mijn promotie.

Mijn vrienden, jullie zijn er nog steeds. Vanwege het drukke leven en mijn ambities heb ik de afgelopen jaren minder aandacht aan jullie besteed dan ik misschien had moeten doen. Ik hoop dat we binnenkort weer wat meer tijd zullen hebben om kroegfilosofie te bedrijven en/of muziek te maken.

Vader, Moeder, bedankt voor de onvoorwaardelijke steun die ik de afgelopen 33 jaar van mijn leven dus ook tijdens mijn promotie van jullie heb gekregen. Jullie hebben ons (Elena en mij) ook praktisch bijgestaan, onder andere door op Christian en Maria te passen. Zij hebben daar klaarblijkelijk van genoten.

Lieve Elena, Ik heb veel van je kunnen leren. Na 10 jaar bij elkaar geweest te zijn, twee kinderen en twee promoties staan wij op een hoogtepunt in ons leven en we hebben elkaar nog meer lief dan ooit. Toen ik begon aan mijn promotie was jij al aan de jouwe begonnen. Ik vroeg je toen of het wel een goed idee was om onze promoties tegelijk te doen. Zoals je toen

al zei, denk ik nu ook, dat als ik er toen niet aan was begonnen, ik er nooit meer aan begonnen was. Ik ben dus blij dat ik het toen wel gedaan heb en dat we het samen hebben kunnen doen. Omdat we in hetzelfde schuitje zaten had je begrip voor wat ik aan het doen was. Je was ook een praktische hulp. Ik weet niet hoeveel mensen kunnen zeggen dat ze hun vrouw raadplegen als ze een moleculair biologisch probleem willen oplossen.

Lieve, lieve Christian en Maria, jullie hebben ontzettend veel geduld opgebracht en mij en mamma onvoorwaardelijk lief gehad. Alle tijd die ik met jullie doorbreng is waardevol. Daarom vind ik dat tijd die ik niet met jullie doorbreng goed besteed moet worden. Ik doe mijn onderzoek, met jullie in mijn hart.

Aernout

List of publications

Yan, Y.; Martens, A.A.; Besseling, N.A.M.; Wolf, F.A. de; Keizer, A. de; Drechsler, M.; Cohen Stuart, M.A. (2008) Nanoribbons self-assembled from triblock peptide polymers and coordination polymers. *Angewandte Chemie-International Edition* 47 (22): 4192 - 4195.

Li, F.; Martens, A.A.; Åslund, A.; Konradsson, P.; Wolf, F.A. de; Cohen Stuart, M.A.; Sudhölter, E.J.R.; Marcelis, A.T.M.; Leermakers, F.A.M. (2008) Formation of nanotapes by co-assembly of triblock peptide copolymers and polythiophenes in aqueous solution. Submitted for publication

Martens, A.A.; Neeleman, M.; Werten, M.W.T.; Vries R.J. de; Eggink, G.; Cohen Stuart, M.A.; Wolf, F.A. de (2008) Triblock protein copolymers forming supramolecular nanofibers and pH-responsive gels at low concentration, irrespective of block order. Submitted for publication

Martens, A.A.; Wolf, F.A. de; Eggink, G.; Cohen Stuart, M.A. (2008) Nucleation and growth by a silk-collagen-like block copolymer into supramolecular nano ribbons. Submitted for publication

Martens, A.A.; Gucht, J. van der; Wolf, F.A. de; Eggink, G.; Cohen Stuart, M.A. (2008) Mechanical properties of silk-collagen-like block copolymer gels. Submitted for publication

Curriculum Vitae

Aernout Anders Martens was born on the 14th of August 1975 in Eindhoven. In June 1994 he graduated from the Hertog Jan College VWO in Valkenswaard, which gave access to an academic education. In September 1994 he started studying Bio Process Engineering at Wageningen University. The study was completed in august 2002 with two master theses and one internship. The first thesis, performed at the Laboratory of Microbiology of Wageningen University, concerned enzymes from the cellulolytic complex of *Clostridium acetobutyricum*. The second thesis and the internship were combined, and concerned hydrogen production by the hyperthermophilic microorganism *Caldicellulosiruptor saccharolyticus* and was performed mainly in the department of Bioconversion at ATO b.v. in cooperation with the Laboratory of Microbiology of Wageningen University. After completing the study, work continued at ATO b.v. as a junior researcher on the subject of hyperthermophilic hydrogen production until May 2003, when a PhD was started under the supervision of Professor Gerrit Eggink, Professor Martien A. Cohen Stuart and Dr. Frits A. de Wolf, on the subject of both the biotechnological production and the physio-chemical properties of polypeptide block copolymers containing charged blocks, the results of which are written in this thesis.

Educational activities

Discipline specific activities

Courses

RPK-C polymer properties	2004
--------------------------	------

Meetings

Student conference (Biezenmortel)	2005
NBC conference (Ede)	2006
Protein congress NWO (Lunteren)	2007
ECIS conference (Geneva)	2007
Student conference (Ven)	2007
Dutch Polymer Days (Lunteren)	2004-2007
Rolduc Polymer Meeting (Kerkrade)	2008

General courses

VLAG PhD week,	2004
Scientific writing, Wageningen University	2006

Optional courses and activities

Preparation PhD research proposal	2003
Weekly FYSKO meeting	2003-2008
Weekly Bioconversuin meeting	2003-2008
Meetings with “DPI gebruikerscommissie”:	2003-2008
Colloquia	2003-2008
Excursion to BASF	2006
PhD trip to Sweden	2007

The research described in this thesis was supported by the Dutch Polymer Institute (DPI) as DPI project 414.

Cover image: Phase contrast microscope image of ammonium sulfate crystals grown in the presence of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$, obtained by drying the $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$ and ammonium sulfate containing solution sampled during the purification of $\text{CS}^{\text{E}}\text{S}^{\text{E}}\text{C}$, after the precipitation steps, but before dialysis.

Cover layout: Elena Martens-Uzunova

Printing: Grafisch bedrijf Ponsen & Looijen b.v. Wageningen