Chapter IV. **Tuning the erosion rate of physically crosslinked artificial protein hydrogels through molecular design and structural control***

Abstract

In this study we demonstrate that the surface erosion rate of hydrogels formed from multi-domain artificial proteins can be tuned through molecular design and structural control. A previously reported protein hydrogel ($AC_{10}A$) assembled through aggregation of leucine zipper domains dissolves at a surface erosion rate of 0.043 mg/cm^{2}min in open solutions. Previous studies of structural and dynamic properties of $AC_{10}A$ hydrogels suggest that the origin of the fast surface erosion is the substantial fraction of intramolecular loops. Here, by harnessing selective molecular recognition of naturally occurring protein motifs, we suppress loop configurations and reduce the surface erosion rate. Specifically, we have designed a triblock protein with two dissimilar coiled-coil domains as the endblocks: the leucine zipper domain A and an α-helical coiled-coil domain (designated P) that constitutes the N-terminal fragment of rat cartilage oligomeric matrix protein (COMP). These two domains do not associate with each other as determined by native electrophoresis of a mixture of the polypeptides P and A, followed by mass analysis of the trypsin digests of the resolved bands. Hydrogels formed from triblock protein $PC_{10}A$ erode 500 times more slowly than $AC_{10}A$ gels and 135 times more slowly than $PC_{10}P$ gels. These three networks have similar stress relaxation times,

suggesting that the differences in erosion rate do not arise from dynamic properties. These gels have substantially different network structures evident in their plateau storage moduli. The plateau storage modulus of a $PC_{10}A$ gel is 5-fold larger than that of an $AC_{10}A$ gel and 1.2-fold larger than that of a $PC_{10}P$ gel at the same pH and concentration, indicating that loops are substantially suppressed in the $PC_{10}A$ network, which leads to slow surface erosion. The role of network structure in regulating material properties is further discussed by a model that relates both modulus and erosion behavior to loop fraction.

1. Introduction

Our group previously reported an artificial protein hydrogel assembled from a multi-domain protein ($AC_{10}A$, sequence shown in Scheme 4.1) through aggregation of the leucine zipper motif A[1]. The capacity for self-assembly is encoded in the protein sequence; therefore, gelation does not require chemical crosslinking reagents, which can compromise material safety in biomedical applications. The modular nature and fidelity of the biosynthetic method used to create these artificial proteins allow different biological determinants—including cell binding domains and enzyme recognition sites—to be readily incorporated in a precisely controlled fashion. These genetically engineered materials have tremendous potential for applications in biomedical fields such as controlled release and tissue engineering. However, as is observed for many other physically crosslinked networks[2-4], $AC_{10}A$ hydrogels dissolve quickly when placed in open solutions. Their rapid dissolution precludes applications where materials are surrounded by excess fluids.

A reduced erosion rate becomes a critical design target for new generations of artificial protein hydrogels. In this study we analyze the origin of the fast erosion of $AC_{10}A$ hydrogels. On the basis of our previous studies on their structural and dynamic properties in closed systems, we believed that suppressing the fraction of looped chains in a network would substantially reduce the erosion rate. Therefore, a new triblock protein $PC_{10}A$ (sequence shown in Scheme 4.1), such that the P domain and the A domain would not associate with each other, was designed, synthesized, assembled and characterized. The P domain (from the N-terminal fragment of rat cartilage oligomeric matrix protein) self-assembles into five-stranded bundles[5,6]. In contrast, the A domain self-assembles into tetramers. We speculated that the different packing structures of P and A should suppress hetero-oligomerization, and that these two domains would not significantly associate with each other.

We experimentally confirm that P and A do not associate with each other and that the resulting suppression of loops in $PC_{10}A$ networks reduces the erosion rate relative to $AC_{10}A$ by orders of magnitude. The differences in material properties of hydrogels formed from different triblock proteins are further discussed on the basis of a model that relates modulus and erosion behavior to the fraction of looped chains.

2. Experimental section

2.1. Construction of expression vectors $pQE9PC_{10}P$, $pQE9PC_{10}A$, and pQE9P

Two DNA segments encoding the P domain were created using the PCR assembly method. Nine oligonucleotides were designed and purchased from Qiagen (Chatsworth, CA). Their sequences are shown as follows.

Sense strands

S-1   5'-GTGAACTGCAGGAAACCAACGCGGCGCTGCAG-3'
S-2   5'-GACGTTCGTGAACTGCTGCGTCAGCAGGTTAAAGAAATCAC-3'
S-3   5'-CTTCCTGAAAAACACCGTTATGGAATCTGACGCGTCTGGTACTAG-3'
S-4   5'-GG ACTAGTGGATCCGGTGACCTGGCGCCGCAGATG-3'

Antisense strands

AS-1   5'-CATAACGGTGTTTTTCAGGAAGGTGATTTCTTTAACCTGCTGAC-3'
AS-2   5'-GCAGCAGTTCACGAACGTCCTGCAGCGCCGC-3'
AS-3   5'-GTTGGTTTCCTGCAGTTCACGCAGCATCTGCGGC-3'
AS-4   5'-GGACTAGTACCAGACGCGTCAGATTC-3'
AS-5   5'-GACG AAGCTTACCAGACGCGTCAGATTC -3'

A dsDNA encoding the P domain flanked by a *Bam*HI site upstream and an *Spe*I

site downstream was assembled from strands S-1, S-2, S-3, S-4, AS-1, AS-2, AS-3, and

AS-4 through two PCR reactions. The sample for the first reaction was prepared as

follows: 0.8 μL of each oligonucleotide solution (ca. 10 μM), 5 μL of 10× reaction

buffer, 1 μL of *PfuTurbo* DNA polymerase, 1 μL of dNTP mix, 2 μL of DMSO, and 36.2

μL of ddH$_2$O. The reaction reagents and buffer were purchased from Stratagene (La

Jolla, CA). The PCR protocol for gene assembly began with initial melting at 95 °C for

30 seconds followed by 10 cycles of denaturation (95 °C for 30 seconds), annealing (52

°C for 30 seconds) and extension (68 °C for 1 minute). The reaction product (1 μL) was

used as the template for the next round of gene amplification with two flanking primers

S-1 and AS-4. The final reaction product was purified with QIAquick Nucleotide

Removal Kit, followed by digestion with *Bam*HI and *Spe*I to generate the overhangs of

these sites. Another dsDNA encoding the P domain flanked by an *Spe*I site and a *Bam*HI

site upstream and a *Hind*III site downstream was assembled from strands S-1, S-2, S-3,

S-4, AS-1, AS-2, AS-3, AS-5 through PCR reactions using the same conditions. This

dsDNA was digested with *Spe*I and *Hind*III to generate the overhangs of these sites. Digestion of pQE9 with *Bam*HI and *Hind*III, followed by ligation with the two P-encoding DNA segments (with the *Bam*HI, *Spe*I overhangs and the *Spe*I, *Hind*III overhangs, respectively) simultaneously, gave the plasmid pQE9PP. This plasmid contains an *Spe*I and a *Bam*HI site between the two P segments. A DNA fragment encoding C_{10} was excised from pQE9C$_{10}$trp with *Nhe*I and *Spe*I, and inserted into the plasmid pQE9PP at the *Spe*I site between the two P segments, yielding the plasmid pQE9PC$_{10}$P.

The plasmid pQE9PC$_{10}$P was digested with *Sph*I and *Hind*III to remove the second P segment and yield linearized pQE9PC$_{10}$. A DNA fragment encoding the A domain was excised from pQE9AC$_{10}$Atrp with *Sph*I and *Hind*III restriction enzymes, and ligated to the linearized pQE9PC$_{10}$ to yield pQE9PC$_{10}$A.

The first P segment was excised out of pQE9PP with *Bam*HI and the remaining plasmid fragment self-ligated to form the pQE9P plasmid.

The sequences of pQE9PC$_{10}$P, pQE9PC$_{10}$A, and pQE9P were verified at the DNA sequencing core facility of the Beckman Institute at the California Institute of Technology.

2.2. Protein expression and purification

pQE9PC$_{10}$P, pQE9PC$_{10}$A, and pQE9P were each transformed into *Escherichia coli* strain SG13009, which carries the repressor plasmid pREP4 (Qiagen, Chatsworth, CA). Proteins PC$_{10}$P, PC$_{10}$A, AC$_{10}$A, P, B[7], and AC$_{10}$ (sequences are shown in Scheme 4.1.) were expressed as described in Chapter II, and purified by affinity chromatography

on a nickel nitrilotriacetic acid resin (Qiagen, Chatsworth, CA) following the denaturing

protocol provided by Qiagen.

2.3. Multi-angle static light scattering

Solutions of the purified protein P were made in 100 mM phosphate buffer (pH

7.6) at concentrations of 30 μM and 107 μM, and subjected to multi-angle static light

scattering measurements on a DAWN EOS light scattering instrument (Wyatt

Technology Corporation, CA). The data were analyzed with Debye plots by using a dn/dc

value of 0.185[8].

2.4. Native electrophoresis, trypsin digestion, and mass spectral analysis

Solutions of AC_{10}, P, and B (200 μM) were prepared in 100 mM phosphate buffer

(pH 7.6), respectively. Solutions of AC_{10}, P, B and their mixtures AC_{10} + B and AC_{10} + P

(at a volumetric ratio of 1:1) were incubated at room temperature overnight. Native

electrophoresis for these solutions was performed on 12% PAGE using the standard

protocol with SDS and reducing agents omitted from all solutions. Solutions of AC_{10}, P,

and B were diluted to 100 μM before being loaded. The gel was stained with Coomassie

Brilliant Blue R-250.

The two protein bands resolved from the AC_{10} + P mixture solution were excised

from the native gel. Coomassie was removed from each protein band[9], followed by

digestion with 0.02 mg/mL trypsin (Promega) at 37 °C overnight. The samples were

purified on a $ZipTip_{C18}$ column (Millipore, Bedford, MA) and eluted with 10 μL of

elution buffer (50% acetonitrile, 50% water, 0.1% trifluoroacetic acid). The MALDI

matrix α-cyano-β-hydroxycinnamic acid (4 μL, 10 mg/mL in 50% CH_3CN) was added to 1 μL of purified protein solution, and 0.5 μL of the mixture was spotted on the sample plate and analyzed on an Applied Biosystems Voyager mass spectrometer.

2.5. Rheological tests

AC$_{10}$A, PC$_{10}$P, and PC$_{10}$A hydrogels (7% w/v) were prepared in 100 mM phosphate buffer at pH 7.6. The gels were centrifuged to remove entrapped bubbles and then loaded on an RFS III rheometer (TA Instruments, New Castle, Delaware). The temperature was controlled at 22.0 $\pm$ 0.1 °C by a Peltier thermoelectric device. A cone-and-plate geometry (0.04 rad cone angle and 25-mm diameter) was used. The edge of the sample was covered with mineral oil to minimize solvent evaporation. Frequency sweep experiments were performed at 1% strain.

2.6. Measurements of erosion rates

AC$_{10}$A, PC$_{10}$P, and PC$_{10}$A hydrogels (6% w/v) were prepared in 100 mM phosphate buffer at pH 7.6, and centrifuged to remove entrapped bubbles. Each gel (60 mg) was transferred into a cylindrical plastic container (8.5 mm diameter and 3 mm height). The container was sealed, and centrifuged at 1700 g for 3 minutes to yield a flat gel film (8.5 mm diameter and ca. 1.06 mm thickness). The gel film, together with the container, was soaked in 3 mL phosphate buffer (100 mM, pH 7.6) in a scintillation vial. The vial was placed on a lab rotator (Lab-Line Instruments, Melrose Park, IL) waving at 60 rpm. Aliquots (80 μl) of supernatant were taken at successive time points and the protein concentrations were determined by measuring the absorbance at 280 nm on a Cary 50 Bio UV-vis spectrophotometer (Varian, Palo Alto, CA). Aliquots were returned

into the vial after each measurement to maintain a constant volume of supernatant. For each $AC_{10}A$ and $PC_{10}P$ gel, eroded fraction at each time point and the erosion rate were calculated on the basis of the absorbance of the solution after the network completely dissolved (indicated by a constant absorbance). Since $PC_{10}A$ gels dissolved slowly (it took ca. two months for test samples to dissolve completely), eroded fraction at each time point and the erosion rate were calculated on the basis of an extinction coefficient of 6970 $M^{-1}cm^{-1}$, which was determined from the amino acid sequence[10].

3. Results and discussion

3.1. Structural and dynamic properties underlying fast erosion of $AC_{10}A$ hydrogels

$AC_{10}A$ hydrogels exhibit extremely fast surface erosion in open solutions. A 1-mm-thick $AC_{10}A$ hydrogel (6% w/v) dissolves completely within 3 hours in 100 mM, pH 7.6 phosphate buffer. Our previous studies on the structural and dynamic properties of $AC_{10}A$ hydrogels in closed systems revealed that these multi-domain protein chains have a strong tendency to form intramolecular loops. The aggregation number of the associative domain (A) is small, and the association is transient in nature. These three factors all contribute to the fast erosion of $AC_{10}A$ networks. Due to intramolecular loops and the small aggregation number of the leucine zipper domain, disengaged clusters form readily (Figure 4.1a). Since the strand exchange time of the leucine zipper domain is on the order of 100 to 1000 seconds near physiological pH, the time scale for disengaged clusters to re-participate into the network is above 100 seconds. In 100 seconds, a cluster with diffusivity of 10^{-7} cm²/s has on average moved 30 μm away from the surface of the network. Therefore, the disengaged clusters generated at the surface of the network are

lost in the surrounding buffer before having a chance to re-connect to the network. The transient nature of the association continuously releases disengaged clusters, leading to a quasi-steady concentration of free clusters at the surface such that their diffusive flux balances their rate of creation at the network surface. Consequently, the network erodes at a constant rate from its surface. The rate of cluster release is governed by the probability of a cluster simultaneously losing all of its connections to the network. We infer that one of the reasons $AC_{10}A$ erodes so quickly is because the fraction of looped chains is substantial. At any given moment, many tetrameric aggregates of the A domain have no connection to the network (Figure 4.1b), and many tetrameric aggregates have configuration such that a single concerted leucine zipper exchange liberates the clusters (Figure 4.1c). Thus, among the structural and dynamic properties that cause the fast erosion of $AC_{10}A$ networks, an essential factor is facile intramolecular association.

3.2. Self-assembling coiled-coil domains P and A do not associate with each other

Fidelity of molecular recognition between protein domains is the basis for many aspects of biological function. Here it provides us the opportunity to control network structure. We believed that intramolecular loops in networks formed from triblock proteins could be substantially suppressed by engineering two dissimilar endblocks that do not associate with each other. In particular, we speculated that the coiled-coil domain derived from the N-terminal fragment of rat cartilage oligomeric matrix protein (COMP)[11] would be likely to prefer homo-oligomerization rather than hetero-oligomerization with leucine zipper A. The coiled-coil domain of COMP associates into five-stranded bundles, while leucine zipper A oligomerizes into tetramers. The different

packing structures of P and A should suppress hetero-oligomerization, and these two domains should not significantly associate with each other.

The gene encoding the coiled-coil domain of COMP was recombinantly constructed into pQE9 plasmid and expressed in *E. coli*. To demonstrate our design concept of tuning erosion rates of purely physically crosslinked networks, the two cysteine residues of the native COMP domain were mutated into serine residues. The protein (designated P) was expressed, and its expected molar mass (6942 Da) was confirmed by mass spectral analysis. Multi-angle light scattering measurements for the 30 μM and 107 μM P solutions revealed average molecular weights of 34310±380 and 35260±160, respectively, suggesting that the cysteine-free domain retains pentameric association (Figure 4.2).

Native electrophoresis of a solution of AC_{10} and P (100 μM, respectively, incubated at pH 7.6 and room temperature overnight) yielded two separate bands in which proteins migrated at the same rates as AC_{10} and P, respectively (Figure 4.3(a)). As a control, retardation in migration of AC_{10} due to its strong association with leucine zipper B was observed on the same gel. Native electrophoresis results suggested that the A domain and the P domain did not associate with each other in the mixture solution. Mass spectral analysis of the trypsin digests of the resolved protein bands provided further confirmation: each band resulted in different digests corresponding to P and A respectively (Figure 4.3(b)).

3.3. Rheological properties of hydrogels assembled from $PC_{10}A$, $PC_{10}P$, and $AC_{10}A$

Triblock proteins $PC_{10}A$ and $PC_{10}P$ were genetically engineered. The molar masses of $PC_{10}A$, $PC_{10}P$, and $AC_{10}A$ proteins are 20860 Da, 20486 Da, and 22105 Da, respectively. In terms of their physical structure, they have almost the same midblocks and their coiled-coil domains have the same length (six heptad repeats). They all assemble into hydrogels in aqueous solutions. Rheological oscillatory shear measurements revealed increased rigidity of the new materials compared to $AC_{10}A$ hydrogels (Figure 4.4). Their G'_∞/nkT values are 0.35±0.01, 0.29±0.02, and 0.07±0.00, respectively, suggesting loops are suppressed in the new materials.

Suppression of looped chains in $PC_{10}A$ gels was expected because the A domain and the P domain do not associate with each other. Decrease in the fraction of looped chains in $PC_{10}P$ gels might result from two structural features of the associative domains. First, the odd aggregation number of the pentameric P domain limits the maximum loop fraction in $PC_{10}P$ networks to be 80%, while there is no such constraint for $AC_{10}A$ networks with tetrameric junctions. Another possible contribution may come from the orientation of the strands in their aggregates. The isolated P domain adopts exclusively parallel orientation in their aggregates[11,12], while our previous studies have shown that the isolated A domain can adopt antiparallel orientation. The length of coiled-coil domains (A or P) with six heptad repeats is 65 Å[13], while end-to-end distances as small as 38 Å for the C_{10} domain were evident in quasi-elastic light scattering measurements (discussed in Chapter II). Therefore, parallel orientation restrains ends of the same molecule from participating in the same aggregate—its midblock has to stretch to form such a loop. Loops form readily in $AC_{10}A$ networks because antiparallel orientation of the A domain

can occur. In contrast, formation of loops in $PC_{10}P$ networks costs energy either to stretch the midblock (if the endgroups are parallel) or to challenge the thermodynamically favored orientation (if the endgroups are antiparallel).

Network relaxation dynamics of $PC_{10}A$, $PC_{10}P$, and $AC_{10}A$ hydrogels are similar. However, as indicated by the position of the peaks in their loss moduli (Figure 4.4), the dominant stress relaxation time of the $PC_{10}P$ gel is noticeably shorter than those of $AC_{10}A$ and $PC_{10}A$ hydrogels. We previously reported that the dominant stress relaxation time of an $AC_{10}A$ hydrogel is strongly correlated with the strand exchange time of the leucine zipper domain. Therefore, network relaxation behavior suggests that the strand exchange rate of the P domain is faster than that of the A domain.

3.4. Erosion behavior of hydrogels assembled from $PC_{10}A$, $PC_{10}P$, and $AC_{10}A$

Despite its faster strand exchange kinetics, introduction of the P domain results in materials of slower erosion rate in open aqueous solutions. The erosion profiles of 6% w/v $AC_{10}A$, $PC_{10}P$, and $PC_{10}A$ hydrogels (Figures 4.5 and 4.6) all show linear mass-loss vs. time profiles, indicating the erosion is occurring at the surface rather than in the bulk. The erosion rates are 4.3×10^{-2} mg/cm^{2}min, 1.3×10^{-2} mg/cm^{2}min, and 9.6×10^{-5} mg/cm^{2}min for $AC_{10}A$, $PC_{10}P$, and $PC_{10}A$ hydrogels, respectively. For 1-mm-thick test samples, $AC_{10}A$ gels dissolved completely within 3 hrs, while $PC_{10}A$ gels dissolved in ca. two months.

Limited stability in open solutions is a common problem for transient networks. For systems such as those formed from hydrophobically modified urethane-ethoxylate (HEUR) polymers, this issue has been addressed by controlling their phase separation

behavior[14-16]. Molecular structure is adjusted to produce a transient network that thermodynamically coexists with its equilibrium sol phase at a low concentration to confer slow surface erosion. The physical mechanism we use here to increase the long-term stability of artificial protein hydrogels in open solutions is different. A constant erosion rate for $PC_{10}A$ gels was observed over a period of 7 days (Figure 4.6) even though the supernatant was not refreshed, indicating that the surrounding solution was far from saturation. Therefore, the slow erosion of $PC_{10}A$ networks is not a consequence of equilibration with the supernatant.

PC$_{10}$A gels erode ca. 500 times more slowly than $AC_{10}A$ gels, and ca. 135 times more slowly than $PC_{10}P$ gels. Since $PC_{10}A$ is constituted from the same associative domains as $AC_{10}A$ and $PC_{10}P$, the significant decrease in erosion rate must not originate from the strand exchange kinetics of the associative domains. The increase in G'_{∞}/nkT for $PC_{10}A$ gels indicates that intramolecular loops are suppressed in these networks. The concomitant decrease in erosion rate supports our original design concept: controlling network structure to suppress loops reduces erosion rates of transient networks formed from artificial proteins.

3.5. A model relating modulus and erosion behavior to loop fraction

Our goal of designing protein hydrogels that erode slowly in open solutions has been accomplished, but several questions remain. How do we reconcile the expected negligible level of intramolecular association in $PC_{10}A$ networks and the deviation of the G'_{∞}/nkT value (0.35) from 1? Why is the difference in storage modulus not as significant

as that in erosion rate? To answer these questions, we offer a model that explains the effects of the fraction of looped chains on the modulus and erosion rate.

Annable proposed a model to predict storage moduli of transient networks formed from telechelic associative polymers[17]. However, this model is not applicable to the systems investigated in this study. Annable's model assumes that micellar functionalities (defined as the number of intermolecular bridges associated with one aggregate) are continuous integral numbers. This is a valid assumption for systems such as those formed from hydrophobically modified urethane-ethoxylate (HEUR) polymers, in which the average aggregation number of the associative group is large (20 to 100)[15] and aggregates are polydisperse. The model does not apply to a system where the aggregation number is uniform and small. Here we derive a model that relates modulus and erosion rate to the fraction of looped chains in transient networks formed from triblock polymers bearing two identical associative endblocks that oligomerize into tetramers ($AC_{10}A$) or pentamers ($PC_{10}P$).

In the solution of a triblock protein bearing two endblocks that associate into tetramers, there are three possible states (Figure 4.7a). They have i=0, 1, and 2 loops, and their functionalities are 4, 2 and 0, respectively. The fraction of aggregates in the state with i loops is designated as f_i. The values of f_0, f_1, and f_2 are related to the overall loop fraction q (the ratio of looped chains to total chains n) through balances on the number of looped chains and the total number of aggregates (n/2):

$$\frac{n}{2}f_1 + 2\cdot\frac{n}{2}f_2 = nq \qquad (1)$$

$$f_0 + f_1 + f_2 = 1 \qquad (2)$$

The relative probability that an aggregate has two loops can be expressed in terms of the relative probability that it has a single loop:

$$f_1 / f_0 \equiv s \tag{3}$$

$$f_2 / f_0 = \frac{1}{2} s^2 \text{ (two loops in state i=2 are indistinguishable)} \tag{4}$$

Solving for s in terms of q:

$$s = \frac{\frac{1}{2} - q - \sqrt{(\frac{1}{2} - q)^2 - 2q(q - 1)}}{q - 1} \tag{5}$$

In terms of s, the fraction of aggregates in each state is:

$$f_0 = 1/(\frac{1}{2} s^2 + s + 1) \tag{6}$$

$$f_1 = s /(\frac{1}{2} s^2 + s + 1) \tag{7}$$

$$f_2 = \frac{1}{2} s^2 /(\frac{1}{2} s^2 + s + 1) \tag{8}$$

If we assume that the probabilities of the occurrence of more complex topologies as illustrated in Figure 4.8 are negligibly small, the modulus and erosion rate of the network to a large extent are determined by the fractions of aggregates in states i=0-2. The ratio of the storage modulus to that of a loop-free network should be close to the fraction of the aggregates in state i=0, because only aggregates of functionalities larger than 2 contribute to network elasticity. Therefore,

$$G'_\infty / G'_{\infty \, (q=0)} \sim f_0 \tag{9}$$

Since the functionality of the aggregates in the more complex topologies illustrated in Figure 4.8 (1) and Figure 4.8 (2) is not larger than 2, neglecting these topologies has no effect on the validity of (9). The erosion rate of the material should be proportional to the fraction of the aggregates in state i=2, which are disengaged from the network and readily lost from the surface.

Similar analysis can be applied for the solution of a triblock protein bearing two endblocks that associate into pentamers, as summarized in Table 4.1. Dependence of f_i on q shows the initially stronger rise in f_1 with q followed by a drop in f_1 as f_2 becomes dominant at large q (Fig. 4.9). For future reference, the values of f_i and $G'_\infty/G'_{\infty\ (q=0)}$ are also tabulated for both tetrameric (Table 4.2) and pentameric (Table 4.3) association.

3.6. Structural properties underlying differences in modulus and erosion behavior of $PC_{10}A$, $PC_{10}P$, and $AC_{10}A$ hydrogels

Here we show that the storage moduli of the gels are less than the values that can be explained by loops alone—failure of the assumptions of the ideal network theory also plays a role. An ideal network assumes incompressibility, affine deformation, no enthalpy change upon deformation, and Gaussian chains[18]. If every molecule in an artificial protein network formed a bridge (q=0, no loops) and the assumptions of the ideal network theory were valid, the modulus would be $G'_{\infty\ (q=0)}$ =nkT. However, rheological measurements show that G'_∞ is substantially less than nkT for all three artificial proteins. This is true even for $PC_{10}A$ gels, which are expected to be nearly loop-free (q≈0). Within the context of ideal network theory, it is not possible to explain $G'_{\infty\ (PC_{10}A)}$/nkT = 0.35. In

the case of $AC_{10}A$ gels, the propensity of protein chains to form loops corresponds to a reduction in the fraction of bridging chains. However, the minimum fraction of bridging chains to produce a network (indicated by an elastic plateau) is the percolation threshold P_c[19], which is 1/3 for tetrafunctional junctions (as in $AC_{10}A$ solutions)[20]. An ideal, tetrafunctional network cannot have a modulus as soft as $G'_{\infty\,(AC_{10}A)}/nkT = 0.07$, because even at P_c the modulus is $G'_{\infty\,(min)}/nkT = 0.15$ (see q=0.67 in Table 4.2).

The artificial protein hydrogels show substantial, systematic deviation from ideal network theory. Qualitatively, this observation accords with what might be expected from the features of the protein chains. One contribution might come from the negative charges in the midblock. Electrostatic repulsion reduces energy cost for the same degree of deformation compared to ideal networks in which only conformational entropy is associated with deformation. Another contribution might come from the fact that the length of the leucine zipper domains and the dimensions of the midblock random coils are comparable. Under strain, the leucine zipper aggregates may rotate, which costs little free energy, to minimize the distortion of the midblock domains.

In this context, it is not clear how to infer the fraction of bridging chains from the observed modulus. Since the $PC_{10}A$ network is expected to be nearly loop-free (q ≈ 0), we use its modulus as a reference modulus $G'_{\infty\,(q=0)} = G'_{\infty\,(PC_{10}A)}$ and it is fairly consistent with the experimental data and the model that we derived. For example, at 7% w/v the modulus of an $AC_{10}A$ gel relating to a $PC_{10}A$ gel is $G'_{\infty\,(AC_{10}A)}/G'_{\infty\,(PC_{10}A)} \approx 0.2$. Given that the gel point of $AC_{10}A$ is ca. 5% w/v, it accords well with $G'_{\infty\,(AC_{10}A)}/G'_{\infty\,(q=0)}$ being a little greater than 0.15 at a concentration just above the percolation threshold and suggests that q≈0.6 in the 7% w/v $AC_{10}A$ network (Table 4.2).

For an $AC_{10}A$ network in which q≈0.6, the fraction of disconnected aggregates f_2≈0.4 (Table 4.2). The erosion rate is proportional to f_2. The high value of f_2, which originates from high q, explains fast erosion of $AC_{10}A$ networks. For an $AC_{10}A$ network to erode 500 times more slowly, f_2 would need to be reduced to ca. 0.001. At that point, q is ca. 0.025 and $G'_\infty/G'_{\infty\ (q=0)}$ is ca. 0.95 (Table 4.2). Thus, the modulus only increases 5 fold. This analysis rationalizes the observation that the effect of loops on erosion rate could be much more significant than its effect on modulus. Since the modulus of a gel must lie between its minimum value at the percolation threshold and its maximum value for q=0, $G'_\infty/G'_{\infty\ (q=0)}$ is constrained to vary less than an order of magnitude.

For a 7% w/v $PC_{10}P$ network, its modulus ($G'_{\infty\ (PC_{10}P)}/G'_{\infty\ (PC_{10}A)} \approx 0.83$) suggests that $q \approx 0.30$ and f_2≈0.18 (Table 4.3). However, an aggregate with i=2 in a pentameric network still has one bridge. The probability that the cluster it bridges to also has only one bridge is f_2, and the probability that its one bridge is to the initial cluster is 1/(number of neighboring aggregates). If the species liberated at the surface is such a dimer, it would track the probability f_2^2/(number of neighbors). This predicts a much greater drop in erosion rate from $AC_{10}A$ to $PC_{10}P$ than is actually observed. Therefore, other free species must contribute to the erosion process. Likely candidates include single aggregates without perfect pentameric association. One possibility is a free species comprising two looped chains. Another possibility is a free species comprising three chains, among which two form loops and the other has a dangling end.

In $PC_{10}A$ networks, if there are not any loops, the smallest object that can be freed is a dimer comprising a tetramer of A domain with all four strands connected to a single cluster of P domain. Therefore, it is quite reasonable that free clusters are much less probable in $PC_{10}A$ networks than in either $AC_{10}A$ or $PC_{10}P$ networks.

4. Conclusions

A design concept of reducing erosion rates of physically crosslinked artificial protein hydrogels through control of network structure was proved. Two coiled-coil domains A and P, which do not associate with each other, were engineered as the endblocks of a triblock protein $PC_{10}A$. $PC_{10}A$ assembled into networks that eroded 500 times more slowly than $AC_{10}A$ networks, and 135 times more slowly than $PC_{10}P$ networks in open aqueous solutions. Similar network relaxation behavior suggests that significant differences in erosion rate do not originate from the strand exchange kinetics of the associative domains. The slow erosion of $PC_{10}A$ networks results from the molecular recognition and self-selectivity of the P and A domains and the consequent low fraction of loops in the networks. A model that relates modulus and erosion behavior to loop fraction was provided to discuss differences in material properties of these hydrogels.

5. References

(1)   Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A. *Science* **1998**, *281*, 389-392.

(2)   Lundberg, D. J.; Brown, R. G.; Glass, J. E.; Eley, R. R. *Langmuir* **1994**, *10*, 3027-3034.

(3)   Kaczmarski, J. P.; Glass, J. E. *Macromolecules* **1993**, *26*, 5149-5156.

(4)   Tam, K. C.; Jenkins, R. D.; Winnik, M. A.; Bassett, D. R. *Macromolecules* **1998**, *31*, 4149-4159.

(5)   Efimov, V. P.; Lustig, A.; Engel, J. *Febs Letters* **1994**, *341*, 54-58.

(6)   Efimov, V. P.; Engel, J.; Malashkevich, V. N. *Proteins-Structure Function and Genetics* **1996**, *24*, 259-262.

(7)   Petka, W. A.; Ph.D. Dissertation; University of Massachusetts Amherst: Amherst, MA 1997.

(8)   Huglin, M. B. *Light scattering from polymer solutions*; Academic Press: London, New York, 1972.

(9)   Hellman, U.; Wernstedt, C.; Gonez, J.; Heldin, C. H. *Analytical Biochemistry* **1995**, *224*, 451-455.

(10)   Gill, S. C.; Vonhippel, P. H. *Analytical Biochemistry* **1989**, *182*, 319-326.

(11)   Malashkevich, V. N.; Kammerer, R. A.; Efimov, V. P.; Schulthess, T.; Engel, J. *Science* **1996**, *274*, 761-765.

(12)   Ozbek, S.; Engel, J.; Stetefeld, J. *Embo Journal* **2002**, *21*, 5960-5968.

(13)   Hodges, R. S. *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **1996**, *74*, 133-154.

(14)   Francois, J.; Maitre, S.; Rawiso, M.; Sarazin, D.; Beinert, G.; Isel, F. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* **1996**, *112*, 251-265.

(15)   Pham, Q. T.; Russel, W. B.; Thibeault, J. C.; Lau, W. *Macromolecules* **1999**, *32*, 2996-3005.

(16)   Tae, G.; Kornfield, J. A.; Hubbell, J. A.; Johannsmann, D.; Hogen-Esch, T. E. *Macromolecules* **2001**, *34*, 6409-6419.

(17)   Annable, T.; Buscall, R.; Ettelaie, R.; Whittlestone, D. *Journal of Rheology* **1993**, *37*, 695-726.

(18)   Flory, P. J. *Principles of polymer chemistry*; Cornell University Press: Ithaca, N.Y., 1953.

(19)   de Gennes, P.-G. *Scaling concepts in polymer physics*; Cornell University Press: Ithaca, N.Y., 1979.

(20)   Albert, R.; Barabasi, A. L. *Reviews of Modern Physics* **2002**, *74*, 47-97.

Scheme 4.1. Amino acid sequences of proteins discussed in Chapter IV.

AC$_{10}$A:
MRGSHHHHHHGSDDDDKASGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEIGDHVAPRDTSYRD
PMG[AGAGAGPEG]$_{10}$ARMPTSGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEIGDHVAPRDTSW

PC$_{10}$A:
MRGSHHHHHHGSGDLAPQMLRELQETNAALQDVRELLRQQVKEITFLKNTVMESDASGTSYRDPMG[AGAGAG
PEG]$_{10}$ARMPTSGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEIGD HVAPRDTSW

PC$_{10}$P:
MRGSHHHHHHGSGDLAPQMLRELQETNAALQDVRELLRQQVKEITFLKNTVMESDASGTSYRDPMG[AGAGAG
PEG]$_{10}$ARMPTSGSGDLAPQMLRELQETNAALQDVRELLRQQVKEITFLKNTVMESDASGKLN

P:
MRGSHHHHHHGSGDLAPQMLRELQETNAALQDVRELLRQQVKEITFLKNTVMESDASGKLN

B:
MRGSHHHHHHGSDDDDKWASGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAKIGDHVAPRDTSM
GGC

AC$_{10}$:
MRGSHHHHHHGSDDDDKASGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEIGDHVAPRDTSYRD
PMG[AGAGAGPEG]$_{10}$ARMPTSW

Abbreviation for domains:
A: SGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAE
P: APQMLRELQETNAALQDVRELLRQQVKEITFLKNTVMESDAS
B: SGDLKNKVAQLKRKVRSLKDKAAELKQEVSRLENEIEDLKAK
C$_{10}$: [AGAGAGPEG]$_{10}$

Table 4.1. Link the fraction of aggregates in each state and the modulus of a network to loop fraction through balances on the number of looped chains and the number of aggregates.

	tetrameric ($\frac{2n}{4}$ aggregates)		Pentameric ($\frac{2n}{5}$ aggregates)	
balance on the number of looped chains	$\frac{n}{2}f_1 + 2 \cdot \frac{n}{2}f_2 = nq$	(1)	$\frac{2n}{5}f_1 + 2 \cdot \frac{2n}{5}f_2 = nq$	(1')
balance on the total number of aggregates	$f_0 + f_1 + f_2 = 1$	(2)	$f_0 + f_1 + f_2 = 1$	(2)
ratio of aggregates in different states	$f_1/f_0 = s$ $f_2/f_0 = \frac{1}{2}s^2$	(3) (4)	$f_1/f_0 = s$ $f_2/f_0 = \frac{1}{2}s^2$	(3) (4)
relationship between the fraction of aggregates in each state and the ratio of looped chains to total chains	$s = \dfrac{\frac{1}{2} - q - \sqrt{(\frac{1}{2}-q)^2 - 2q(q-1)}}{q-1}$ $f_0 = 1/(\frac{1}{2}s^2 + s + 1)$ $f_1 = s/(\frac{1}{2}s^2 + s + 1)$ $f_2 = \frac{1}{2}s^2/(\frac{1}{2}s^2 + s + 1)$	(5) (6) (7) (8)	$s = \dfrac{\frac{2}{5} - q - \sqrt{(\frac{2}{5}-q)^2 - 2q(q-\frac{4}{5})}}{q-\frac{4}{5}}$ $f_0 = 1/(\frac{1}{2}s^2 + s + 1)$ $f_1 = s/(\frac{1}{2}s^2 + s + 1)$ $f_2 = \frac{1}{2}s^2/(\frac{1}{2}s^2 + s + 1)$	(5') (6) (7) (8)
Storage modulus	$G'_\infty/G'_{\infty\,(q=0)} \sim f_0$	(9)	$G'_\infty/G'_{\infty\,(q=0)} \sim f_0 + f_1$	(9')

n	total chains
q	ratio of looped chains to total chains
f_i	fraction of the aggregates having i loops
$G'_\infty/G'_{\infty\,(q=0)}$	normalized plateau storage modulus
s	constant

Table 4.2. Fractions of tetrameric aggregates in states i=0-2 and the normalized storage modulus of the network at different loop fractions

q	f_0	f_1	f_2	$G'_\infty/G'_{\infty\,(q=0)}$
0.1	0.8169	0.16619	0.0169	0.8169
0.2	0.65969	0.28062	0.05969	0.65969
0.3	0.52177	0.35647	0.12177	0.52177
0.4	0.4	0.4	0.2	0.4
0.5	0.29289	0.41421	0.29289	0.29289
0.6	0.2	0.4	0.4	0.2
0.7	0.12177	0.35647	0.52177	0.12177
0.8	0.05969	0.28062	0.65969	0.05969
0.9	0.0169	0.16619	0.8169	0.0169

Table 4.3. Fractions of pentameric aggregates in states i=0-2 and the normalized storage modulus of the network at different loop fractions

q	f_0	f_1	f_2	$G'_\infty/G'_{\infty\,(q=0)}$
0.1	0.77552	0.19896	0.02552	0.97448
0.2	0.58856	0.32288	0.08856	0.91144
0.3	0.42903	0.39194	0.17903	0.82097
0.4	0.29289	0.41421	0.29289	0.70711
0.5	0.17903	0.39194	0.42903	0.57097
0.6	0.08856	0.32288	0.58856	0.41144
0.7	0.02552	0.19896	0.77552	0.22448
0.8	0	0	1	0

(a)

In 100 s, the cluster has on average moved 30 μm from the surface

The time scale for it to re-connect to the network is above 100 s

(b)

(c)
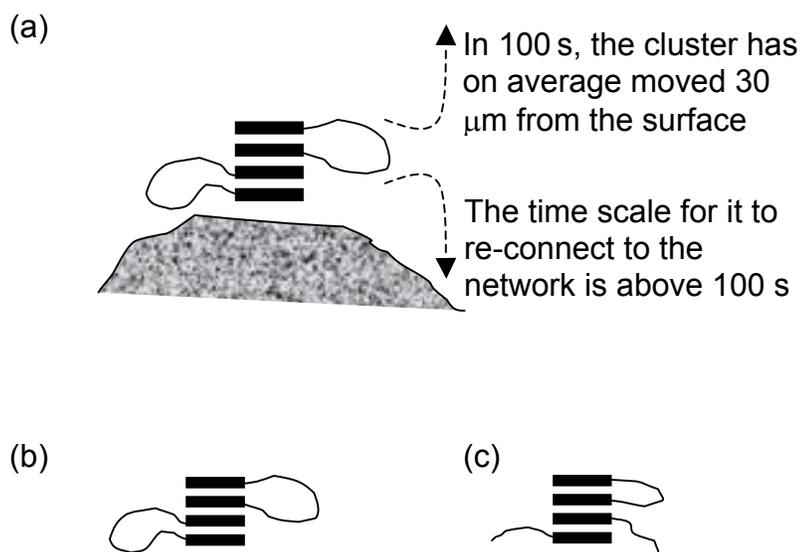
Figure 4.1. Structural and dynamic properties underlying the fast erosion of $AC_{10}A$ hydrogels. (a) Disengaged clusters form readily in the system because of the strong tendency of intramolecular association and the small aggregation number of the associative domain. They are lost from the surface through diffusion before having a chance to re-connect to the network, which takes much longer time due to the slow strand exchange rate. (b) A tetrameric aggregate of the A domain having no connection to the network. (c) A tetrameric aggregate of the A domain having configuration such that a single concerted leucine zipper exchange liberates the cluster.
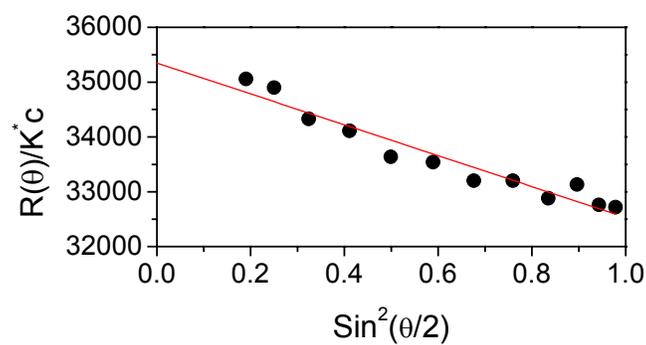
Figure 4.2. Debye plot of multi-angle light scattering signals from a 107 μM solution of P (100 mM phosphate buffer, pH 7.6, room temperature) reveals that the molecular weight of the dominant species is 35260±160, suggesting this cysteine-free domain remains pentameric association (the molecular weight of monomer P is 6942)
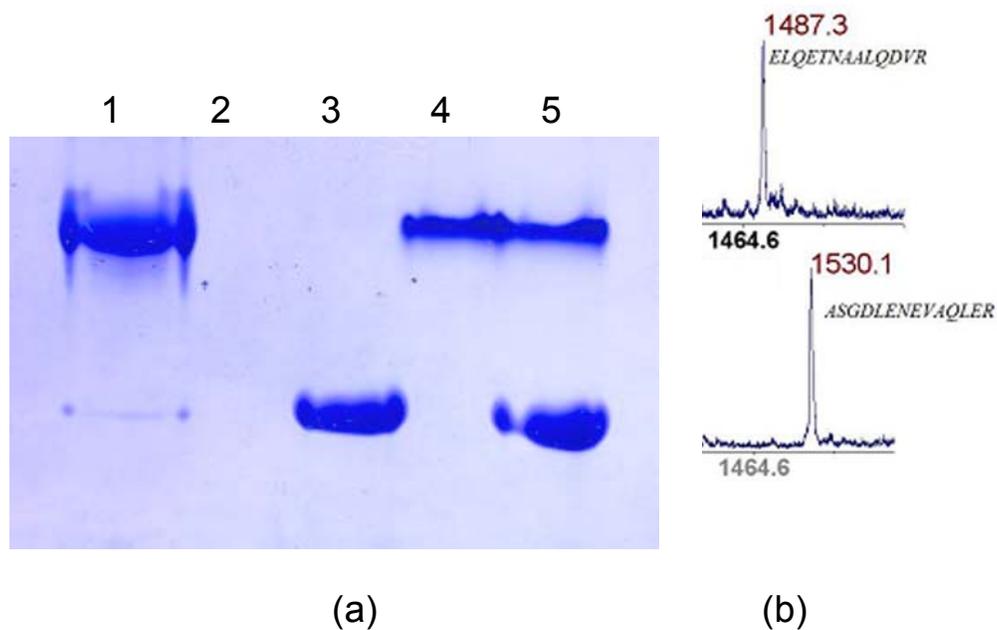
(a)   (b)

Figure 4.3. Coiled-coil domains A and P do not associate with each other. (a) Native electrophoresis of recombinant proteins. Lane 1. AC_{10} + B; 2. B; 3. AC_{10}; 4. P; 5. AC_{10} + P. Protein B by itself does not migrate into the gel due to its net positive charges. (b) Mass spectral analysis of trypsin digests of the proteins in the two bands excised from lane 5.
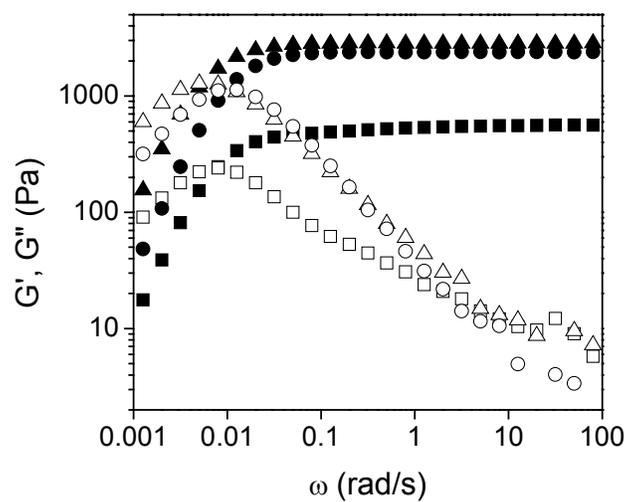
Figure 4.4. Dynamic moduli (closed symbols for storage modulus; open symbols for loss modulus) of $AC_{10}A$ (■, □); $PC_{10}P$ (●, ○); and $PC_{10}A$ (▲, △) hydrogels. (7% w/v, 100 mM phosphate buffer, pH 7.6, 22 °C)
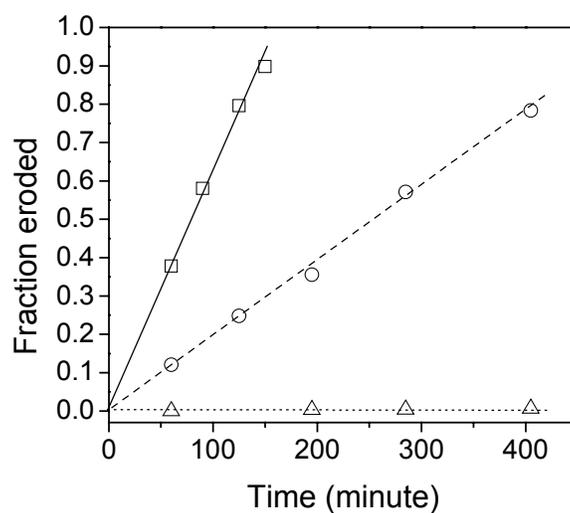
Figure 4.5. Erosion profiles of AC$_{10}$A ($\square$); PC$_{10}$P ($\bigcirc$); and PC$_{10}$A ($\triangle$) hydrogels. (6% w/v, 100 mM phosphate buffer, pH 7.6, room temperature) The surface area of each gel is 0.5672 cm^2. The total mass of each gel is 60 mg. The erosion rate: 4.3×10^{-2} mg/cm^{2}min for AC$_{10}$A and 1.3×10^{-2} mg/cm^{2}min for PC$_{10}$P.
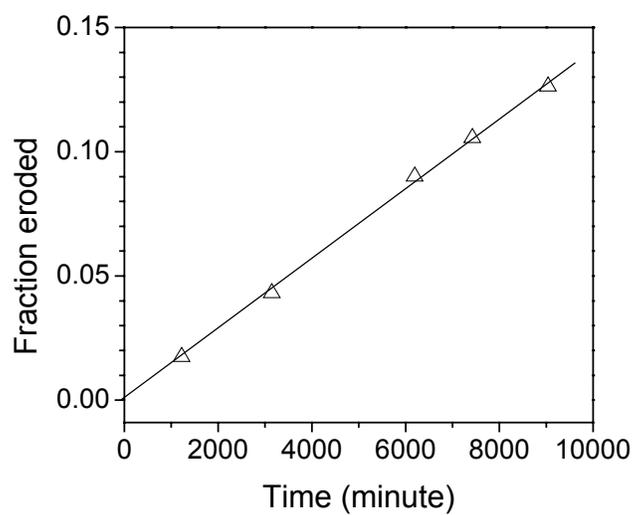
Figure 4.6. The erosion profile of a PC$_{10}$A hydrogel. The surface area of the gel is 0.5672 cm^2. The total mass of the gel is 60 mg. The erosion rate is 9.6×10^{-5} mg/cm^{2}min. (6% w/v, 100 mM phosphate buffer, pH 7.6, room temperature)
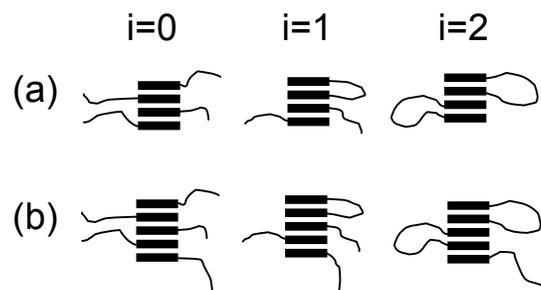
i=0   i=1   i=2

(a)

(b)

Figure 4.7. Three possible states of aggregates designated by the number of loops, i, in each. (a) tetrameric aggregates, the functionality for i=0, 1, and 2 is 4, 2, and 0, respectively. (b) pentameric aggregates, the functionality for i=0, 1, and 2 is 5, 3, and 1, respectively.
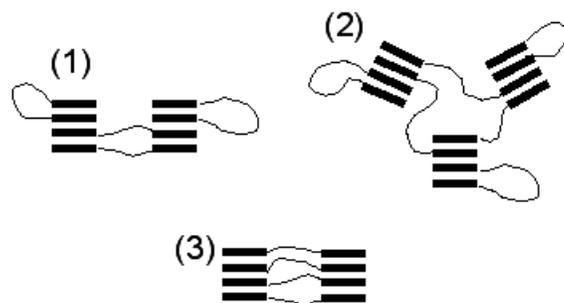
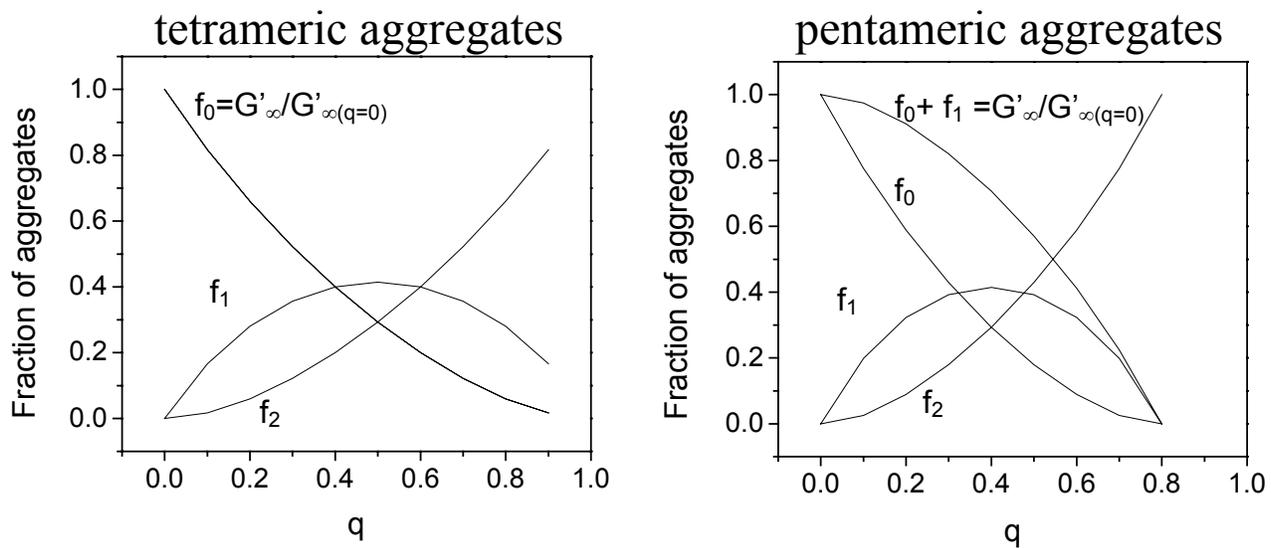Figure 4.8. More complex topologies in systems containing tetrameric aggregates.

Figure 4.9. Fractions of aggregates in different states as functions of the fraction of

loops.