2 Background and Review of Relevant Literature

2.1 Introduction

The purpose of this chapter is to introduce the main principles relevant to the presented research. These principles are essential to understanding the goal of producing surface-active poly(amino acids) (or PAA's) that self-assemble on metal oxide surfaces to form brush layers. In addition, this review aims to provide some of the motivation behind the research objectives.

This chapter is broken up into two major sections. Section 2.2 introduces the basic concepts of surface-active polymers and brush formation. This discussion begins with a brief overview of the theory behind polymer adsorption and brush formation. Next, one of the most widely used applications of brush-forming polymers (steric stabilization of colloidal suspensions) is introduced. The section closes with a brief discussion of the disadvantages of conventional polymer synthesis techniques and the potential of biosynthetically-derived PAA's as surface-active brush forming polymers.

Subsequent sections (Sections 2.3-2.5) cover the concepts of producing biosynthetically-derived PAA's. These sections include brief overviews of recombinant DNA technology, protein expression, purification methods and biochemical characterization techniques. Section 2.3.2 includes a discussion on the difficulties of applying current cloning strategies to the production of surface-active PAA's that will provide some proof that the new strategies proposed here are of crucial importance.

2.2 Surface-Active Polymers

2.2.1 Polymer Adsorption

Surface-active polymers are polymers that have been designed with the intention to interact specifically with a particular surface. The main type of interaction is known as *adsorption*. There have been many studies investigating the process of polymer adsorption onto solid surfaces. One of the most extensive studies in this area has been conducted by Fleer and Scheutjens.[1] This theory employs a self-consistent-field lattice model where the space adjacent to the surface is discretized into layers of lattice sites.[1] This approximation allows for the infinite number of polymer chain conformations in a

given space to be reduced to a finite number. The resulting theory is able to predict polymer adsorption and its dependence on several parameters. These parameters include polymer molecular weight and bulk concentration, the Flory polymer-solvent interaction parameter (χ) and the energy of adsorption per segment ($\chi_s kT$, where subscript *s* denotes that the parameter refers to the polymer segment). The segmental adsorption energy is important since it dictates whether a polymer will adsorb or not. As $\chi_s kT$ increases, the tendency to adsorb also increases. Adsorption occurs when the polymer monomer possesses a greater affinity for the surface than the solvent does. Hence, segmental adsorption energy is an important parameter to consider when designing a polymer to interact with a specific surface.

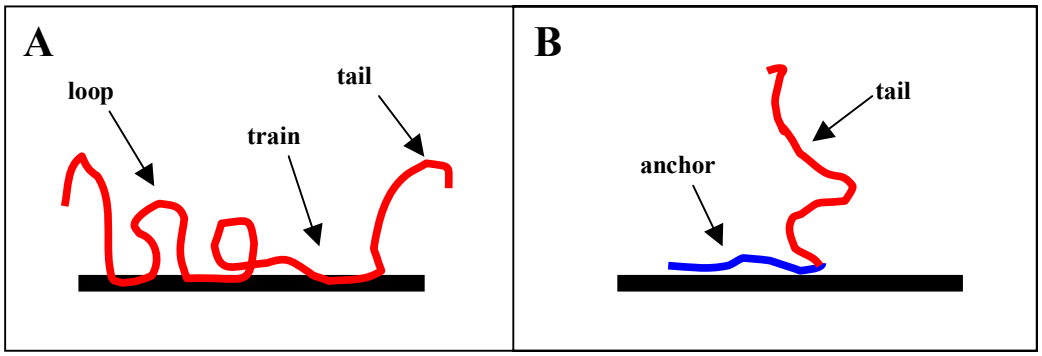

Figure 2.1 (A) 'Loops, Trains and Tails' conformation of a homopolymer and (B) block copolymer adsorption with brush formation.

2.2.2 Brush Formation

During adsorption on a surface, polymers form various conformations. Two general cases are shown in Figure 2.1. Figure 2.1A shows adsorbed homopolymers forming 'loops, trains and tails'. Figure 2.1B illustrates the adsorption of a diblock copolymer, consisting of a short adsorbing block and a longer block that extends away from the surface into the surrounding medium, forming a brush layer.

Surface-active polymers that self-assemble into brush layers are usually block copolymers. Block copolymers are composed of at least two distinct moieties. The *anchor* block preferentially adsorbs to the target surface. The interaction between the anchor and the surface may be ionic or specific in nature. The *tail* block has a high affinity for the solvent and forms the brush layer by extending away from the surface. Extensive studies have been conducted on the theory of block copolymer adsorption and the conformation of the resulting adsorbed layers.[2,3,4,5] It is required that the tail block

does not adsorb onto the surface and is completely soluble in the dispersing medium ($\chi < 0.5$) . Therefore, in terms of adsorption interaction energies, $\chi_{s,anchor}kT > \chi_{s,tail}kT$, where $\chi_{s,tail}kT$ should be as small as possible, preferably = 0, and χ_s is a dimensionless segmental adsorption energy parameter.[1]

As indicated above, the first step in the design of a brush-forming polymer is to choose the proper anchor/tail blocks for a particular solid/solvent system. The next important design parameter is the polymer composition. There are several theories that have been developed to determine the optimum ratio of the number of tail repeat units to the number of anchor repeat units. The scaling theory developed by Marques and Johanny predicts the optimum ratio for the case of block copolymer adsorption in a nonselective solvent (where both blocks are soluble).[6] Marques and Johanny proposed that for optimum tail extension and maximum polymer adsorption, the brush layer thickness, δ, scales as:[6]

Equation 2.1 Scaling theory prediction of adsorbed layer thickness:

$$\delta \sim \left(\frac{N_{tail}}{N_{anchor}^{1/3}} \right)$$

where N_i = degree of polymerization of block *i*

In Equation 2.1, it can be seen that the brush thickness of the adsorbed layer scales linearly with the degree of polymerization (or molecular mass) of the tail block (N_{tail}).

Marques and Johanny's scaling theory agrees relatively well with other theories based on the self-consistent field theory using a lattice model.[1] The results proposed by these theories have been also been verified by experimental work done by several groups. Wu, Yokomama and Setterquist studied the effect of anchor and tail block sizes on the copolymer adsorption on SiO_2 in 2-propanol.[7] This group found that the *optimum anchor block composition is ~5mol%* (or % number of repeat units). This agrees well with Marques and Johanny's scaling theory. These results serve as a useful guide when designing a block copolymer that will form optimal brush layers.

Some additional examples in the literature of block copolymers that demonstrate the ability to self-assemble on solid surfaces to form brush layers are: poly(vinylmethyl ether)/poly(vinyloxy-4-butryic acid) copolymers on α-Fe_2O_3 in water,[8] poly((2-

dimethylamino)ethyl methacrylate)/poly(sodium methacrylate)) copolymers on TiO_2 in water,[9] poly((2-dimethylamino)ethyl methacrylate)/poly(n-butyl methacrylate) copolymers on SiO_2 in isopropanol[7] and poly-(L-lysine)/poly(ethylene glycol) copolymers on net negatively charged biological surfaces in aqueous environments.[10]

2.2.3 Applications of Surface-Active Polymers

There are many applications where surface-active polymers that self-assemble to form brush layers are used. They are used in a variety of areas such as microelectronics, adhesives, coatings, cosmetics, pharmaceuticals and drug delivery.[11,12,13] Recently, there has been interest in using self-assembling polymers that display a variety of different biomolecules. The goal of these works is to design new modified surfaces that can be used for analytical biodevices[14,15,16, 17], bioseparations[18,19], biocompatible surfaces[20] and for the investigation of protein folding pathways[21,22]. For each application, appropriate block copolymers are designed to interact with a particular surface.

Brush-forming block copolymers are widely used as dispersants in the processing of colloidal suspensions. Particles are classified as *colloids* when they have at least one dimension that is on the order of a micron. We have chosen to design our surface-active polymer to form brush layers on aluminum oxide surfaces for the purpose of colloidal stabilization. Colloidal alumina suspensions are important for industrial applications, such as in the preparation of paints, the treatment of pigments and ceramic powder processing.[23] Section 2.2.3.1 discusses briefly the theory behind colloidal suspensions and how brush-forming polymers are used in this field.

2.2.3.1 Steric Stabilization of Colloidal Metal Oxide Suspensions

When dealing with colloidal suspensions, it is of great importance to understand the forces that act upon the particles. The most widely accepted theory to describe these interactions is known as DLVO theory. This theory assumes that the more long-ranged interparticle interactions mainly control colloidal stability.[24] Two types of forces are mainly considered: *van der Waals* and *electrostatic*. According to DLVO theory, the total interaction potential (V_T) is the sum of the attractive van der Waals forces (V_A) and repulsive electrostatic double-layer forces (V_R), as shown in Equation 2.2:[25]

Equation 2.2 DLVO Expression for the Total Interaction Potential (V_T):

$$V_T(D) = V_R(D) + V_A(D)$$

where D = minimum separation distance between particle surfaces

This relationship can be best explained through the use of the 'potential energy – distance curve' shown in Figure 2.2A. Positive values of V_T correspond to repulsive forces and negative values correspond to attractive forces. To obtain the curve for the total potential energy of interaction, the curves for V_A and V_R are added together.

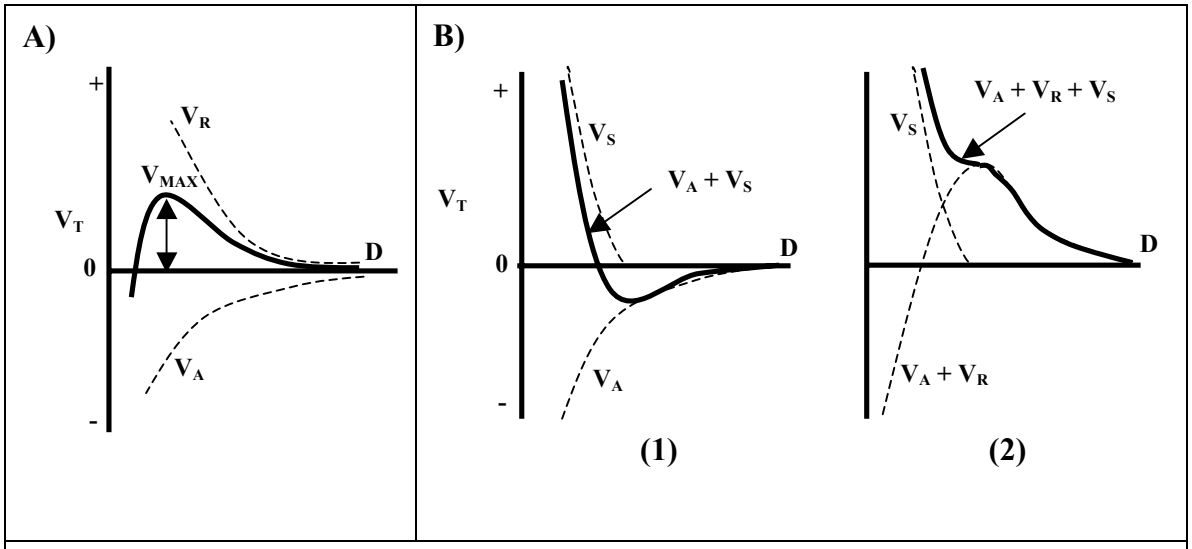

Figure 2.2 Potential energy-distance curves. (A) Total interaction potential energy between two spheres; (B) Interaction energy curves for sterically stabilized particles: (1) in the absence of electric double layer repulsion ($V_T=V_A+V_S$); (2) in the presence of an electric double layer repulsion ($V_T=V_A+V_R+V_S$). (Figures adapted from Ref. 26; reprinted from Ref. 26 with permission from Elvesier).

In Figure 2.2A, V_A and V_R are plotted individually to show their dependence on D. The van der Waals attraction dominates at small and large values of D. However, if sufficiently large enough, electrostatic repulsive forces may predominate at intermediate interparticle distances. This can lead to a potential energy barrier with a magnitude denoted as V_{max}.[27] If the potential energy maximum is large compared to the thermal energy of the system ($\geq 10kT$) the suspension is kinetically stabilized for periods of months to years.[28] When $V_{max} \leq 1$kT, then particles can approach one another to distances where van der Waals forces will cause them to aggregate in periods of seconds to minutes, resulting in an unstable suspension.

DLVO theory provides the basic understanding for the complex system of forces acting upon colloidal particles. It dictates that the potential energy barrier to aggregation must be maximized in order to impart stability. Steric stabilization is one method of increasing this potential energy barrier. During steric stabilization, large molecules are adsorbed onto the colloid surfaces that *sterically* hinder the particles from closely approaching one another. These molecules are usually polymeric in nature and may be nonionic or ionic. By modifying DLVO theory to include the repulsive potential supplied by steric forces (V_S) , Equation 2.2 becomes:[24]

Equation 2.3 DLVO Expression for the Total Interaction Potential (V_T) including Steric Term:

$$V_T(D) = V_A(D) + V_R(D) + V_S(D)$$

Figure 2.2B shows the total potential energy curves when a steric force is present. As shown, two cases may arise. In the case of nonionic polymeric stabilizers, the only repulsive contribution is due to steric forces (1). Ionic polymeric stabilizers contribute both steric and electrostatic repulsions (2). In each case, there is an infinite energy barrier that prevents the particles from aggregating and forming an unstable suspension (see Figure 2.3).[29]

When colloidal alumina suspensions are rendered stable by steric stabilization, the suspensions become easier and cheaper to process. Steric stabilizers reduce processing costs because viscosities are dramatically lowered, therefore enabling the processing of suspensions with high solids content.[30,31] Therefore, the development and synthesis of effective polymeric brush-forming stabilizers has become an important area of research.

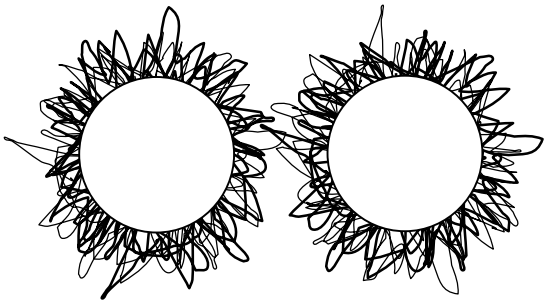

Figure 2.3 Sterically stabilized colloidal particles. (Adapted from Ref. 32; reprinted from Ref. 32 with permission from Elvesier).

2.2.4 Conventional Synthetic Polymer Production

There are currently many techniques to synthesize conventional block copolymers. Some of these techniques include free radical polymerization[33] and living polymerization reactions.[34,35] These techniques continue to be improved and optimized. However, due to the nature of the polymerization reactions, it is difficult to synthesize a monodisperse polymer with a defined monomer sequence.[34] In addition to the lack of control, conventional polymerization reactions also require numerous complicated experimental steps that usually involve toxic monomer reagents and harmful organic solvents.

In recent years, there has been a trend to develop polymers that are non-toxic and biodegradable.[36,37] Protein-based polymers, or polypeptides, offer this advantage. Proteins are made up of 20 naturally occurring amino acid monomers. These amino acids differ by their side chain. These side chains span the range of physio-chemical properties: hydrophilic/hydrophobic, hydrogen-bond donors/acceptors, acidic/basic, small/bulky. It is the primary sequence arrangement of these 20 amino acid monomers that leads polypeptides to self-assemble into a multitude of complex structures, such as α-helices, β-strands and random coils.[38] These stable and highly ordered structures dictate the unique 3-dimensional structure and, ultimately, the biological function of each protein.

Much research has been done trying to determine how these highly ordered structures form. It has been hypothesized that by applying these concepts to the design of synthetic polypeptides, new synthetic or bio-inspired materials can be produced.[39] In theory, by specifying the amino acid sequence, one can design a synthetic PAA to have a defined secondary structure. The initial attempts to synthesize PAA's of defined composition and structure used solid phase peptide synthesis. However, these synthesis reactions are limited to polypeptides of <50 amino acids because they are plagued by unwanted side reactions and a lack of control over the molecular weight distribution.[40]

Deming demonstrated a novel synthesis technique for making high molecular weight block copolypeptides with well-defined sequences using the ring-opening polymerization of α-amino acid-N-carboxyanhydride (NCA) monomers and organo-nickel compounds as initiators.[40] Although the synthesis reactions Deming proposed were an improvement over conventional organic polymer synthesis techniques, they still have their disadvantages. The polymerization reactions involve reagents that are not readily

available, organic solvents and highly controlled reaction conditions. Using this method, synthetic polypeptides were constructed one amino acid at a time, through a series of reactions. However, this method still results in a polydisperse polymer product (M_w/M_n = 1.13). This method is also limited in the fact that it can only be used to synthesize block copolymers. For example, this technique could not be used if one wanted to vary the monomer sequence within a block. It is apparent that the conventional block copolymer synthesis methods presently lack the level of control that is necessary to produce high molecular weight polymers that have a well-defined monomer sequence and composition.

2.3 Biosynthetically-derived PAA's

2.3.1 Biosynthesis vs. Chemical Synthesis

For the facile synthesis of high molecular weight synthetic polypeptides with a well-defined amino acid sequence and composition, we decided to turn to the protein synthesis machinery of living organisms. Using recombinant DNA technology, the gene sequence encoding for a defined synthetic polypeptide sequence can be constructed and cloned into a genetically engineered organism. Using this genetic information, the organism can express the synthetic polypeptide as if it was a native protein. McGrath *et al.* compared using this biosynthetic method to chemical synthesis methods.[41] This work showed that the biosynthetic approach to producing block copolypeptides offers better efficiency and a higher level of control over monomer sequence and molecular weight. These recent advances in molecular bioengineering and polymer chemistry indicate that interest the area of bioinspired polymeric materials is rapidly growing.[39]

The objective of the work presented here is to show the feasibility of producing highly precise, biosynthetically-derived, surface-active PAA's. Using a newly developed cloning strategy to analyze the effectiveness of this approach, surface-active PAA's were expressed within *E. coli*, purified and biochemically characterized.

2.3.2 Recombinant DNA Technology

Over the last 10 years, the field of recombinant DNA technology has increased exponentially. The ability to manipulate DNA by researchers has facilitated significant scientific achievements, such as the sequencing the complete genomes of several organisms and the re-engineering of proteins to explore structure/function relationships. Section 2.3.2.1 introduces some conventional cloning techniques and Section 2.3.2.2 reviews some of the literature covering recent advances in the biosynthetic production of PAA's. Both of these sections include the disadvantages of these techniques and provide evidence for the necessity of the new cloning strategy developed in this work.

2.3.2.1 Conventional methods

Several significant developments have been made that have lead to the exponential growth of the field of recombinant DNA technology. These developments include the chemical synthesis of oligonucleotides, the polymerase chain reaction (PCR) and restriction enzyme (RE) cloning. These technologies shape the conventional methods used in the cloning of synthetic genes.

Using these methods, a synthetic gene can be constructed, amplified and inserted into a plasmid vector for gene expression. The first step of this process is accomplished by chemically synthesizing oligonucleotides with a defined nucleotide sequence that encodes for a specified amino acid sequence. Automated synthesizers, commonly employing the beta-cyanoethyl phosphoramidite method, make these synthetic oligonucleotides readily available and relatively inexpensive. By applying hybridizing technologies, single-stranded oligonucleotides can be used to construct short double-stranded DNA molecules (Figure 2.4A). Using PCR, these genes can be amplified to sufficient amounts for subsequent cloning steps. PCR employs a thermostable DNA polymerase (*Taq* DNA polymerase) and short synthetic oligonucleotides, called primers, to hybridize to and amplify a target DNA sequence (Figure 2.4B). Restriction enzymes can then be used to clone the PCR product into a recipient plasmid vector. A plasmid is a circular, self-replicating DNA molecule (Figure 2.4C). Restriction enzymes recognize short nucleotide sequences and cleave the DNA near that site. Over the years, many restriction enzymes have been identified and isolated for these types of applications. For

efficient cloning, restriction enzymes that leave overhangs, or 'sticky ends', are used. These overhangs are then used to ligate the DNA molecule into a recipient plasmid that has complimentary overhangs. The information within this DNA segment then allows for the desired polypeptide to be expressed.

Though the method discussed above has been used for many years, it has some significant disadvantages, especially for the construction of the long genes required to produce PAA's that contain repeating sequences. First, the process of making synthetic oligonucleotides longer than 150 bases is currently very difficult. Oligonucleotides are synthesized one base at a time through a series of reactions, each of which can be controlled to be ~99% efficient. However, since the efficiency is less than 100%, the overall percent yield of the oligonucleotide synthesis decreases with increasing oligonucleotide length. Eventually, the yield decreases to the point that it becomes difficult to control product quality. A few vendors state they can custom make oligonucleotides up to 200 bases long, but the high cost of these services makes this an unattractive option. Therefore, if a long gene is required, a method for piecing together smaller DNA molecules is required. Also, when using restriction enzymes for cloning, sometimes extra base pairs need to be included in the DNA sequence for enzyme recognition sites. These extra bases can lead to unwanted amino acids in the final PAA product. Therefore, for the production of PAA's, conventional cloning methods are insufficient and improved methods are necessary.

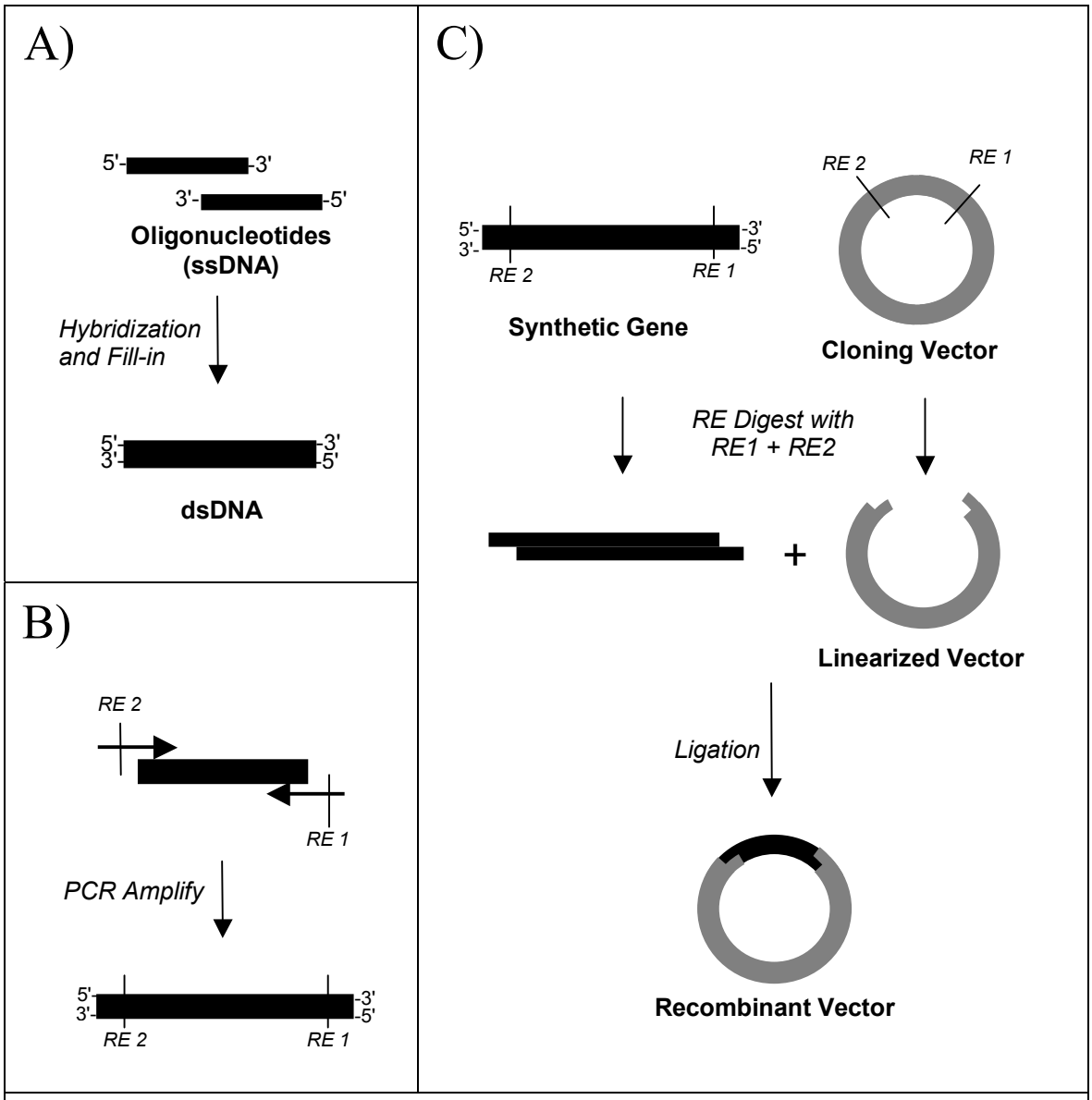

Figure 2.4 Schematic illustrating conventional synthetic gene assembly. (A) Hybridization of single-stranded synthetic oligonucleotides to create double stranded DNA molecule. (B) PCR amplification using synthetic primers (arrows) to add restriction enzyme (RE) recognition sites. (C) RE enzyme cloning of synthetic gene into a linearized vector to construct a recombinant vector.

2.3.2.2 Prior Cloning Strategies for the Production of PAA's

Initial Work (Cappello J et al.)

Initial improvements in the ability to construct long repeats of short DNA sequences for the production of PAA's came with the strategic use of special restriction enzymes. Cappello *et al.* were one of the first groups to examine the feasibility of producing synthetic protein polymers using recombinant DNA technology.[42] Cappello *et*

al. were able to construct genes encoding for silk-like and elastin-like polymers from short monomer segments by using a series of restriction enzymes. Their strategy is outlined in Figure 2.5.

They began by cloning a small DNA 'monomer' encoding for 6 repeats of a 6-amino acid sequence into a cloning vector using the restriction enzyme *Ban*II (Figure 2.5A). The vector was then propagated within *E. coli* and isolated. The monomer was released from the vector using *Ban*II and self-ligated. The resulting 'multimers' (made up of 2 or more monomer units) were then ligated back into a cloning vector (Figure 2.5B). Again, the plasmid was propagated and isolated. The multimer gene was then released using a different restriction enzyme, *Ban*I, and cloned into an expression vector (Figure 2.5C). Nagarsekar *et al.* were able to use this method to produce a gene containing 16 repeats of a 138 bp monomer, resulting in a silk-elastinlike copolymer of ~66 kDa.[43]

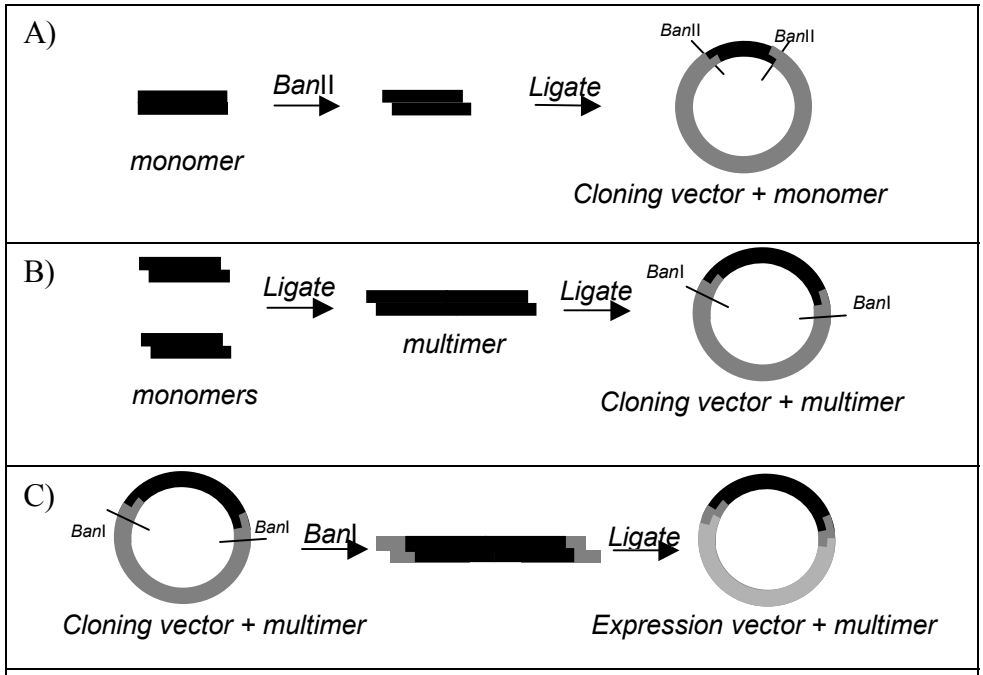

Figure 2.5 Schematic illustrating the cloning method developed by Cappello *et al*.

This cloning technique has several drawbacks, which are discussed below and summarized in Table 2.1. First, the strategy includes the incorporation of restriction enzyme recognition sites into the gene which ultimately add unwanted amino acids to the final protein polymer. The final PAA produced has an extra 78 bp due to restriction enzyme cloning, encoding for an extra 26 amino acids. Secondly, the strategy uses

restriction digestion to release gene segments from plasmids, which can be a slow and inefficient technique, as well as expensive depending on the enzyme used. Thirdly, this technique requires the use of a cloning vector in addition to an expression vector. It would be more desirable, for convenience purposes, to be able to do the cloning steps within the expression vector. Lastly, the self-ligation of gene monomers is an unpredictable process (as will be discussed later) resulting in less control over gene length as well as unwanted ligation products. The cloning strategy developed by Cappello *et al.* is important in that it demonstrated the feasibility of producing protein polymers using recombinant DNA technology.

Seamless Cloning (Padgett KA and Sorge JA)

Padgett and Sorge contributed to this area of research with their development of a 'seamless cloning' strategy. Padgett and Sorge's strategy took advantage of PCR and the type-IIS restriction enzyme, *Eam*1104I.[44] This enzyme has the useful ability to cut outside its DNA recognition sequence (5' ctcttcn|nnn 3', where n is any of the four bases) and can also be inhibited by site-specific methylation. Seamless cloning begins with the PCR amplification of a gene with primers that contain *Eam*1104I sites. When the PCR product is digested with *Eam*1104I, 'sticky ends' are created. This is accomplished without leaving behind any evidence of the enzyme's recognition sites. If there is a recognition site within the gene, PCR can be conducted in the presence of 5-methyldeoxycytosine (^{m5}dCTP). This incorporates a methylated cytosine into the internal recognition sites, which prevents the enzyme from cutting at those locations. In addition, the cloning vector is linearized by PCR in the presence of ^{m5}dCTP using primers that contain *Eam*1104I sites. The PCR-amplified vector is then digested, leaving 'sticky ends'. The gene is then ligated into the linear plasmid, leaving no evidence of the *Eam*1104I sites. Hence, the insertion is referred to as *seamless*. Padgett and Sorge's process is shown in Figure 2.6.

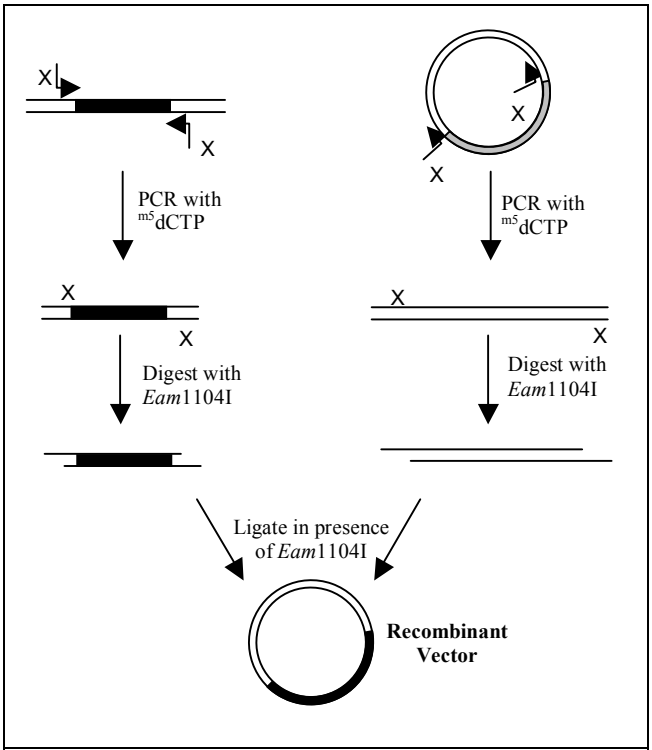

Figure 2.6 Schematic diagram of the seamless cloning method. X represents *Eam*1104I recognition sites introduced by PCR primers (Adapted from Ref. 44; reprinted from Ref. 44 with permission from Elvesier).

Seamless cloning provides the ability to clone genes into plasmid vectors without the incorporation of recognition sequences that affect the final amino acid sequence. One of the disadvantages of this method is that ^{m5}dCTP incorporation during PCR can be inefficient. Therefore, if there are many internal recognition sites, it becomes increasingly difficult to incorporate an ^{m5}dCTP at each site, resulting in unwanted restriction fragments after digestion with the enzyme. However, the seamless cloning technique is quite useful and has been applied in the construction of PAA's, as will be discussed.

Applying Seamless Cloning Method (McMillan RA et al. and Goeden-Wood NL et al.)

McMillan *et al.* applied the seamless cloning technique developed by Padgett and Sorge to show the feasibility of rapidly assembling synthetic genes encoding for protein polymers.[45] McMillan *et al.* were able to construct long repeats of short DNA sequences encoding for an elastin-like protein. This was accomplished by PCR amplifying a short DNA monomer with primers containing *Eam*1104I sites in the presence of ^{m5}dCTP. The PCR product was purified and digested, followed by a self-ligation step. Multimers of the monomer were then ligated into a linearized expression plasmid. Ultimately, multimers of 3000 bp were obtained, 40 repeats of a 78 bp monomer. Expression analysis showed that a 90 kDa elastin-like PAA was produced using this method.

Goeden-Wood *et al.* described a similar method with a few improvements.[46] In this work, a 432 bp gene consisting of 9 repeats of a 48 bp DNA monomer was constructed and used to produce a 17 kDa protein polymer designed to have a β-sheet

structure. The improvements made were aimed at solving the problem of inefficient ^{m5}dCTP incorporation during PCR. To accomplish this, a new vector was constructed to eliminate internal recognition sites by site-directed mutagenesis. In addition, a more stringent restriction enzyme, *Sap*I, was used. This enzyme is similar to *Eam*1104I but requires an extra base in its recognition site (5' gctcttcn|nnn 3'), thus reducing the probability of multiple recognition sites within the vector.[46] These improvements increased the efficiency with which the multimeric gene could be constructed.

The work of the McMillan and Goeden-Wood research groups is significant in the area of biosynthetic PAA production. Both groups were able to successfully produce a highly-ordered synthetic structural polymer by applying novel recombinant DNA technologies. However, some limitations exist (Table 2.1). The methods utilized rely on the self-ligation reaction to 'polymerize' the gene monomers. As mentioned, this reaction is difficult to control. As shown in the work of each group, the size of the final polymer could not be controlled. This may be acceptable for some polymer applications and characterization studies, however, for most applications a well-defined polymer size is crucial (*e.g.* brush-forming applications).

Recursive Directional Ligation (Meyer DE and Chilkoti A)

Meyer and Chilkoti developed a stepwise approach to ligate short DNA monomers seamlessly in a procedure they termed as 'recursive directional ligation' (RDL) in order to produce an elastin-like PAA.[47] Meyer and Chilkoti developed this method due to the inefficiency of self-ligation reactions. They stated that although the self-ligation of DNA monomers with cohesive ends was rapid, it sacrificed precise control because it is a statistical process that results in a population of DNA multimers with a distribution of different lengths.[47]

The RDL cloning strategy is illustrated in Figure 2.7 and summarized in Table 2.1. As shown, a short DNA monomer is constructed and inserted into a cloning vector using restriction enzyme cloning. The monomer is designed to have two different restriction enzyme recognition sites (RE1 and RE2) on either end. Therefore, the plasmid can be subjected to two reactions.

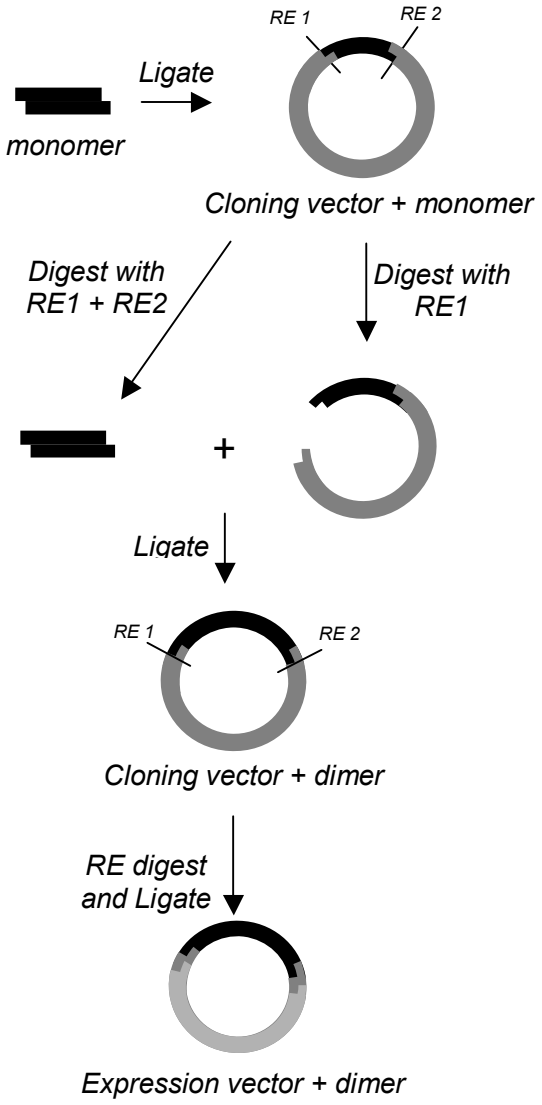

Figure 2.7 Schematic diagram of the RDL cloning method. (Adapted from Ref. 47; reproduced from Ref. 47 with permission from American Chemical Society).

One reaction consists of a digestion by both RE1 and RE2 to free the monomer gene. The other reaction consists of the linearization of the plasmid by digesting with only RE1. The free monomer can then be ligated back into the linear vector, resulting in a 'dimer' gene. The sites for RE1 and RE2 are regenerated at both ends of the gene with no internal sites. Therefore, the process can be repeated until a gene of the desired length is achieved. The final multimeric gene is then freed from the cloning vector using additional restriction enzymes and subsequently cloned into an expression vector.

Using a DNA monomer encoding for 50 amino acids (~150 bp), Meyer and Chilkoti were able to construct a gene consisting of up to 30 repeats encoding for a PAA of ~130 kDa. This method allows for the precise control over multimer length, unlike previous methods. However, the use of two restriction enzymes is undesirable. As stated by Meyer and Chilkoti, there are stringent requirements that need to be met during the selection of RE1 and RE2. The enzymes chosen must have different and unique recognition sites, produce complimentary ends and be compatible with the coding sequence of the gene.[47] Unfortunately, there is only a limited pool of restriction enzymes. Therefore, this method is not universally compatible with any desired DNA sequence. In

addition, as mentioned, restriction enzyme cloning and the use of separate cloning and expression vectors are also undesirable characteristics.

Controlled Cloning (Won JI and Barron AE)

Won and Barron definitively showed that using the head-to-tail self-ligation method of multimerizing short DNA monomers is an inefficient way of creating long, repetitive DNA sequences.[48] Using the seamless cloning method developed by McMillan *et al.*, Won and Barron were unable to achieve DNA multimers longer than 6 repeats of a 63 bp DNA monomer. They discovered that this was due to *circularization* of the multimers once they reached a particular length. This process is shown in Figure 2.8. As a result, the long DNA multimers were unavailable for ligation into a recipient plasmid. Won and Barron determined that the seamless cloning method was insufficient for producing long repetitive DNA sequences.

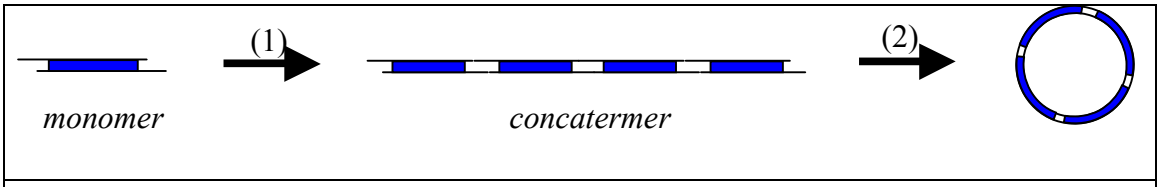

Figure 2.8. Schematic illustrating the circularization of DNA concatemers. (1) monomers forming concatemers and (2) their subsequent circularization that occurs before they can be cloned into a recipient vector.

Won and Barron subsequently developed a new cloning technique for this purpose. Their technique also involved the use of *Eam*1104I and *Sap*I and the inability of each to cut at recognition sequences containing ^{m5}dCTP. DNA monomers were constructed so that they contained an *Eam*1104I site at the 5' end and a *Sap*I site at the 3' end. The monomer was then divided into two reactions: one digested with *Eam*1104I and the other with *Sap*I. The monomer digested with *Eam*1104I was left with 5' overhangs on each end. The monomer digested with *Sap*I only had one 5'overhang. Next, the *Eam*1104I-digested monomer was dephosphorylated with calf intestinal alkaline phosphatase (CIP) so that the monomers could not self-ligate. Then, both digested monomers were ligated together to form a 'dimer'. The dimer was re-phosphorylated using T4 polynucleotide kinase (T4 PNK), and ligated into a linearized recipient plasmid, which was constructed in the same manner as in the seamless cloning method. The dimer

gene was propagated and then amplified from the plasmid using PCR with primers that re-established the 5' *Eam*1104I and 3' *Sap*I sites so that the process could be repeated until a multimer of desired length is achieved. This process is shown in Figure 2.9.

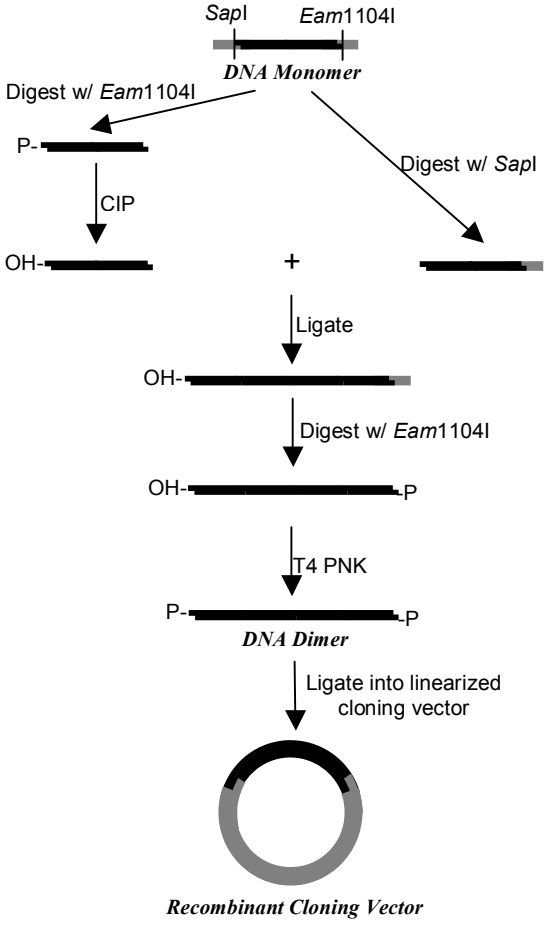

Figure 2.9 Schematic diagram of the Won and Barron method. (Adapted from Ref. 48; reproduced from Ref. 48 with permission from American Chemical Society).

With this technique, Won and Barron were able to construct multimeric genes with 6, 12, 24 and 48 DNA monomer repeats encoding for PAA's of 151, 277, 529 and 1033 amino acids in length, respectively. The approach developed by Won and Barron takes advantage of the seamless cloning method but also enables the precise control of the size of the multimerized gene. However, there are some limitations to this method (Table 2.1). The full incorporation of ^{m5}dCTP into an expression vector during PCR is inefficient, so it is difficult to remove internal restriction sites. In addition, two different plasmid vectors for cloning and expression are required for this method. It is advantageous to use only one vector for both cloning and expression. This method requires numerous experimental steps where it would be advantageous to have a more rapid cloning procedure.

The cloning techniques discussed above have made tremendous progress towards the ability of biosynthetically producing PAA's. However, there are still some limitations that need to be addressed in order to make this type of technology useful for efficient biosynthesis of brush-forming PAA's. The new cloning strategy developed in this

research moves towards making these improvements and making the design and production of synthetic PAA's more routine.

Method	Key Features	Disadvantages
J. Cappello *et al.*	• Pioneering cloning strategy for producing synthetic protein polymers • Uses *Ban*II to create small DNA monomers with 'sticky ends' • Ligation of monomers to form multimers • Uses *Ban*I to free multimer gene from cloning vector and insert it into expression vector	• RE sites add unwanted amino acids in final PAA sequence • Monomers obtained through inefficient RE digest • Self-ligation of monomers - circularization • Separate cloning and expression vectors
K.A. Padgett and J.A. Sorge	• "Seamless Cloning" • Relies on PCR and unique RE (*Eam*1104I) to create small DNA monomers • Ligation of monomers to form multimers • No unwanted amino acids in final sequence due to RE's • After multimerization, gene is RE cloned from cloning vector into expression vector	• Relies on ^{m5}dCTP incorporation into vector during PCR • Monomers obtained through inefficient RE digest • Self-ligation of monomers - circularization • Separate cloning and expression vectors
D.E. Meyer and A. Chilkoti	• "Recursive Directional Ligation" • Relies on RE's that are compatible with desired sequence for 'seamless' cloning • Step-wise approach to build multimers from small DNA monomers • More control over gene length • Uses separate RE's to free gene from cloning vector and insert it into expression vector	• Limited number of compatible RE's (not universal for any desired sequence) • Monomers obtained through inefficient RE digest • RE's used to clone gene into expression vector may add unwanted amino acids in final sequence • Separate cloning and expression vectors • Numerous experimental steps
J.I. Won and A.E. Barron	• Uses *Eam*1104I and *Sap*I to improve control over self-ligation reaction of monomers – avoids circularization of long multimers • Step-wise approach to build multimers from small DNA monomers • No unwanted amino acids in final sequence due to RE's • After multimerization, gene is RE cloned from cloning vector into expression vector	• Relies on ^{m5}dCTP incorporation into vector during PCR • Monomers obtained through inefficient RE digest • Separate cloning and expression vectors • Numerous experimental steps

Table 2.1 Summary of cloning techniques discussed, including key features and the disadvantages of each. (RE = restriction enzyme).

2.4 Recombinant Protein Expression

Using recombinant DNA techniques, including those discussed in Section 2.3.2, many advances have been made in the design and production of recombinant proteins and protein-based polymers. This section covers some of the aspects of recombinant protein expression, such as choice of expression system and the incorporation, and subsequent removal, of fusion tags that simplify downstream purification. These aspects will be used in the production of biosynthetically-derived, surface-active PAA's.

2.4.1 Host Organism

The first step in designing the production of a recombinant protein is the choice of host organism. Current recombinant DNA technologies allow recombinant proteins to be expressed in a variety of organisms from bacteria to yeast to mammalian cells. The choice is made depending on the desired properties of the final protein. Bacterial systems are well understood and their genetics are easily manipulated. However, if extensive post-translational modifications are required for protein activity, then a more complex host organism may be required.[49]

For the production of surface-active PAA's, bacterial systems are ideal, since they are relatively simple to work with and can express large amounts of protein in a relatively short amount of time. *Escherichia coli* is the chosen host in most cases based on its convenient culture characteristics, well understood gene expression systems and ability to produce large amounts of recombinant protein. Many studies have also been done with *E. coli* to optimize recombinant protein production, such as genetically improving upon the protein expression yield by using strong promoters or by the use of high-usage codons.[50,51] It is apparent that *E. coli* would be the ideal host for the facile production of surface-active PAA's.

2.4.1.1 Bacterial Expression Vectors

There are many prokaryotic expression systems that are commercially available for the production of recombinant proteins. These systems usually consist of self-replicating, circular DNA molecules, called *plasmids*. Each system has its own unique characteristics that can be tailored towards a variety of different applications.

When choosing an appropriate expression vector for the production of a particular recombinant protein, several characteristics must be considered. Most vectors will express the protein of interest as a fusion protein to aid in expression efficiency and purification. If a fusion tag is desired, the type of tag, whether the tag is C- or N-terminal to the protein of interest and the method of tag removal must be considered.

The method in which the gene encoding for the protein of interest is inserted into the vector is also an important decision. For example, this may be accomplished using multiple cloning sites (MCS), which consists of various sites for restriction enzyme cloning. The recently developed method of TOPO® TA cloning allows for the rapid insertion of PCR products. This decision is usually based upon the flexibility one has with the DNA sequence encoding for the protein of interest.

The desired level of recombinant protein expression and the location of the final product must also be considered. Cells can be programmed to overexpress the protein of interest or tightly regulate expression of proteins that may be toxic to the host strain. The location of the final protein product can be chosen to be intracellular or extracellular. For example, an expression system using *Caulobacter crescentus* can be used to secrete the recombinant protein into the medium to simplify purification by incorporating a secretion signal peptide.[52]

Additional considerations include the choice of antibiotic selection and the presence of an epitope for immunodetection of the recombinant protein. By determining the characteristics desirable for a particular experiment, an appropriate and effective expression system can be chosen. If an appropriate commercially available expression vector can not be found for a particular application, modifications can be made. In the presented work, a commercially available vector has been modified to make it satisfy the specified requirements so that it can serve as both a cloning and an expression vector using the newly developed cloning strategy.

2.4.2 Affinity Fusion Tags

The use of affinity fusion tags in recent years has been a powerful tool in recombinant protein production. Affinity fusion tags are polypeptide sequences that are 'fused' with the protein of interest. The tag simplifies production of the protein of interest

by providing a known "handle" that can be used in the purification and detection of the recombinant protein.

Currently, there are several different affinity fusion tags being used for recombinant protein production and purification. Some of the most common are: polyhistidine tags,[53,54] Histidine-Patch Thioredoxin,[55] maltose-binding protein (MBP),[56,57] chitin-binding domain (CBD)[58] and glutathione S-transferase (GST) fusions.[59] Each fusion system incorporates a specific interaction that can be exploited in order to simplify purification by chromatography. Polyhistidine tags and Histidine-Patch Thioredoxin have an affinity for resins charged with immobilized divalent metal ions (Ni^{+2}, Cu^{+2}). MBP, CBD and GST fusions have an affinity for amylose resins, chitin beads and resins containing immobilized glutathione, respectively.

Affinity fusion tags can also increase expression levels, promote solubility and offer an epitope for immunodetection. One such example is the Histidine-Patch Thioredoxin fusion tag system. This tag is a recombinant version of the native thioredoxin protein which acts as a hydrogen donor for ribonuclease reductase.[60] This recombinant form was created by switching two specific amino acids to histidine, resulting in the formation of a 'histidine patch' on the surface of the protein.[61] The Histidine-Patch (HP) has a very high affinity for divalent cations, a property which can be exploited in order to purify fusion proteins on metal-chelating resins. HP-Thioredoxin also increases the solubility of the fusion protein, making it less likely that the recombinant protein will form inclusion bodies.[62] This is beneficial since inclusion body formation can complicate downstream purification. According to the manufacturer, HP-Thioredoxin fusions are able to accumulate to approximately 40% of the total cellular protein. In addition to aiding purification, HP-Thioredoxin can also serve as an epitope for immunodetection using commercially available antibodies.

In the presented research, polyhistidine and HP-Thioredoxin affinity fusion tags were used to aid in the production of surface-active PAA's. These fusion tags allow for the purification of the PAA's from the soluble cell extract using immobilized metal affinity chromatography (IMAC). The cell extract, which contains many of proteins in addition to our PAA fusion, is loaded onto a chromatography column containing a metal chelating resin (*e.g.* iminodiacetic acid (IDA)-Sepharose) charged with Ni^{+2}. The affinity

fusion tags bind to the immobilized nickel ion, while unbound proteins are washed away. Weakly bound proteins are eluted by washing in a buffer of lower pH. Finally, the target protein, the HP-Thioredoxin fusion, is eluted by lowering the pH further or adding imidazole which competes for the immobilized metal ion sites. This purification method is quite simple and effective and has been used in the purification of polyhistidine- and HP-Thioredoxin-affinity tagged proteins such as D-lactate dehydrogenase[63] and human tissue transglutaminase.[64]

2.4.3 Fusion Tag Removal

In most cases, it is desirable to remove the affinity tag once the fusion protein has been purified. The methods for tag removal fall into two main categories: enzymatic and chemical. Ideally, the removal method should be specific, mild and simple. Specificity is needed to ensure that only the affinity tag is removed from the protein of interest. Mild conditions are required so that the protein is not denatured or destroyed, which is especially important when protein activity is a concern. Lastly, simple and rapid removal methods are preferred so that the overall protein production process is not further complicated.

2.4.3.1 Enzymatic Digestion

Enzymatic removal of the fusion tag is accomplished by engineering a protease recognition site between the affinity tag and the protein of interest. Protease enzymes recognize specific amino acid sequences and cleave the protein at a defined site. Enterokinase is an example of a serine protease that is commonly used for the removal of N-terminal fusion tags. A serine protease binds to a substrate by recognizing a specific short amino acid sequence. The resulting complex places the active site of the enzyme near the cleavage site. The active site of a serine protease contains a serine and a histidine within close proximity of one another. The side chains of these amino acids interact with the substrate, ultimately leading to the breaking of a peptide bind. Enterokinase recognizes the amino acid sequence $(Asp)_4LysX$ (where X is any amino acid) and hydrolyzes the peptide bond between the Lys and X residues.

Enzymatic removal of fusion tags has several limitations. Protease enzymes can be inefficient at times and very costly. For example, the efficiency of enterokinase can be reduced by the amino acid downstream of the lysine in the recognition site to 60% (when X=Pro).[65] Lavallie *et al.* developed a recombinant version of enterokinase, which consists of just the catalytic subunit, to increase the specific activity.[66] However, this alternative is more expensive, which must be taken into account during production scale-up. In addition, a further purification step is required to remove the free affinity tag and the protease enzyme. Thus, enzymatic affinity tag removal is useful, but not ideal.

2.4.3.2 Chemical Cleavage

Chemical cleavage is a rapidly growing method for fusion tag removal. These methods include using cyanogen bromide to cleave the bond to the C-terminal side of Met residues,[67] low pH-induced cleavage at Asp-Pro bonds[68] and using hydroxylamine to cleave at Asn-Gly bonds.[69] With relatively mild reaction conditions of pH 9.0 and 45°C, hydroxylamine chemical cleavage is an attractive choice. Also, it is more specific than the other chemical cleavage methods mentioned above. Bornstein and Balian reported that the Asn-Gly pairing is a relatively uncommon one (~0.25% of amino acid pairs),[69] which suggests that relatively few proteins sequences will have internal cleavage sites.

The hydroxylamine cleavage reaction is shown in Figure 2.10. The reaction begins with the cyclization of the side chain of the asparagine to form a cyclic imide. This is followed by the nucleophilic addition of two hydroxylamine molecules to the cyclic imide. The resulting intermediates rearrange, leading to the liberation of a new polypeptide with an N-terminal glycine residue. Hence, an affinity tag can be released from the protein of interest by incorporating an Asn-Gly sequence between the two domains. The only requirement is that the final protein product will have an N-terminal glycine residue or C-terminal asparagine residue.

This method has been applied in the production of several recombinant proteins such as rhIGF-I[70] and an antimicrobial peptide.[71] The reactions can be relatively quick (~4 hrs) and the reagents are inexpensive. To stop the reaction, the hydroxylamine can be dialyzed away against any desired buffer. However, chemical cleavage does have its

disadvantages. Minor side reactions have been reported to occur that result in cleavage of Asn-Leu, Asn-Met, and Asn-Ala resides[69] and hydroxamates of Asn and Gln residues.[72]

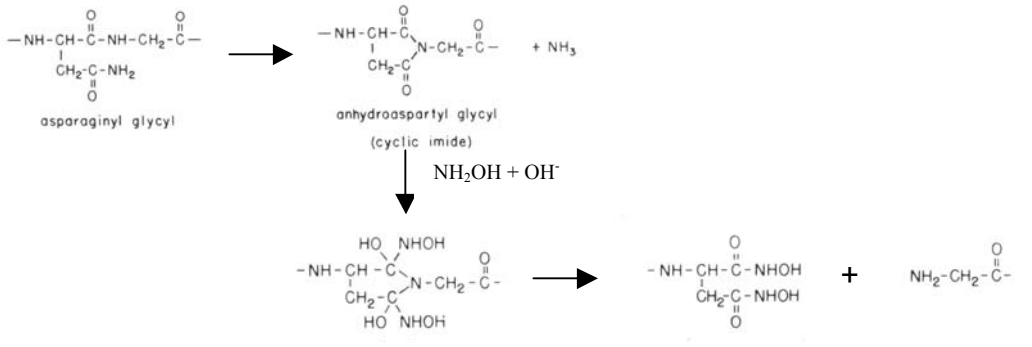

Figure 2.10 Hydroxylamine cleavage reaction. (Adapted from Ref. 69; reprinted from Ref. 69 with permission from Elvesier).

2.5 Protein Techniques

After the surface-active PAA is expressed and its fusion tag is removed, several commonly used protein techniques are employed in order to isolate and verify the identity of product. SDS-PAGE, HPLC, amino acid compositional analysis and MALDI-TOF MS were used to characterize the unique PAA's produced. The purpose was to confirm the amino acid composition and molecular weight of the products, thus validating the research presented by demonstrating that the designed PAA's were in fact produced and isolated. Each method is briefly discussed below.

2.5.1 SDS-PAGE

Sodium **d**odecyl **s**ulfate – **p**oly**a**crylamide **g**el **e**lectrophoresis (SDS-PAGE) is one of the most commonly used techniques for analyzing proteins. This method allows for the relatively rapid determination of the molecular masses of constituent proteins in a sample. Its use has become so widespread due to the fact that it is relatively simple, requires only nanogram to microgram amounts of protein, and its ability to process multiple samples simultaneously.

The electrophoretic separation of proteins is based upon the phenomenon that a molecule with a net charge, in this case a protein, will migrate when subjected to an

electric field. In most cases, electrophoretic separations are carried out in gels. Polyacrylamide gels are used here since they are chemically inert and stable over a wide range of conditions (*e.g.* pH, temperature).[73] In addition, the gel pore size can be controlled through the changing of the concentrations of monomer (acrylamide) and crosslinking agents (bisacrylamide). The pore size can be controlled so that the gel acts as a molecular sieve, where small proteins will travel faster through the porous network than larger proteins.

During SDS-PAGE, proteins are denatured and therefore, the separation is based mostly on the basis of their molecular weight. Before the separation takes place, the protein sample is mixed in a buffer solution containing SDS. SDS is an anionic detergent that will disrupt the structure of a native protein and cause it to denature. In other words, the SDS breaks up the 3-dimensional shape of the protein, resulting in a linear, rod-like structure. SDS binds to most polypeptides with a ratio of 1 dodecyl sulfate molecule per 2 amino acid residues.[74] The SDS-protein complex has a large net negative charge that is roughly proportional to the mass of the protein.[75] Therefore, most proteins will have the same charge-to-mass ratio. The protein solution is then loaded onto a polyacrylamide gel and subjected to an electric field. The negatively charged SDS-protein complexes migrate through the polyacrylamide gel towards the positive cathode and are separated only by their size. The proteins are then visualized by various staining methods (*e.g.* Coomassie or silver stain) and their molecular weights can be determined by comparison with a standard which contains a series of proteins of known molecular weight.

After electrophoresis, the proteins must be stained in order to visualize the separation. There are several staining techniques, where Coomassie Blue R-250 and silver staining are the most common. Coomassie Blue R-250 staining is a simple procedure where the gel is incubated in a methanolic solution containing the dye. The dye stains the proteins bands resulting in a bluish color. Once excess dye is washed away, the background is clear. Coomassie Blue R-250 can detect as little as 0.1 μg protein. However, different proteins will stain with different intensities and the exact mechanism of the staining is unknown. It has been shown that the dye strongly binds to hydrophobic and positively charged proteins, as well as proteins rich in arginine, lysine and histidine.[76]

Silver staining, introduced in 1979 by Switzer *et al.*,[77] is more sensitive (~1 ng) but it is a more laborious procedure consisting of numerous washing steps. The mechanism of silver staining is the reduction of silver nitrate to metallic silver at a protein which has been fixed in the gel using glutaraldehyde. This leads to the deposition of silver grains at the location of a protein and a band of brownish color. Once again, different proteins stain differently with silver and the precise mechanism of silver staining is unknown.

By varying factors such as the pore size, buffer conditions and addition of reducing agents, SDS-PAGE can be tailored to almost any application for the separation, detection and quantification of proteins. SDS-PAGE can also be used as a first step before immunodetection (*i.e.* Western blotting). Therefore, SDS-PAGE is an extremely useful technique for analyzing proteins because it is simple, rapid, sensitive (1 ng – 0.1 μg) and can resolve proteins that differ in mass by ~2%. It will be used throughout the work presented here to monitor protein expression and purification.

2.5.2 HPLC

High **p**erformance **l**iquid **c**hromatography (HPLC) has become a valuable tool in the field of biotechnology where it has been used to identify, purify and characterize complex biomolecules.[78] The HPLC separation of a mixture of solutes (*e.g.* peptides or proteins) arises from the differential adsorption of each solute according to their varying affinities for the immobilized stationary phase.[79] HPLC offers several modes of separation, where the most commonly used are reversed-phase (RP) and ion-exchange (IEC). Each method provides the ability to separate closely related molecules with excellent resolution. HPLC also has the advantages of speed (minutes not hours), sensitivity (femto- to nanograms), reproducibility (1%), high recovery, accuracy and the ability to be automated.[78]

The key component of HPLC is the stationary phase, which is usually in the form of derivatized silica or polymer-based particles (d=3-10 μm) packed into a stainless steel column.[80] The basis of separation is determined by the type of derivitization used. For example, for RP-HPLC, the particles are derivitized in order to immobilize n-alkyl ligands on their surface. Thus separation is mediated through hydrophobic interactions

between the solute and these ligands.[79] These n-alkyl ligands can range from butyl (C4) to octadecyl (C18) functional groups. For ion exchange HPLC, binding is mediated through electrostatic interactions between a charged solute (*e.g.* protein) and charged immobilized functional groups.[79] The immobilized functional groups commonly used for anion exchange HPLC are diethylaminoethanol (DEAE) and quaternary aminoethyl (QAE). For cation exchange, carboxymethyl (CM) and sulfopropyl (SP) are used.

In addition to the stationary phase, there are several other factors that affect the separation of molecules. Most of these factors are determined by the choice of mobile phase, which is the phase containing the solute that is being pumped through the column. For RP-HPLC, the main considerations for choosing an appropriate mobile phase are the selection of pH, ion-pairing agents and organic modifiers. Peptide and protein separations can be sensitive to the pH of the mobile phase, due to the different ionizable groups within these molecules. Ion-pairing agents, like TFA, are often used to neutralize the ionic groups of the solute and increase its hydrophobicity.[78] Organic modifiers are used to partition the solute between the stationary and mobile phase. Separation of the solutes is often accomplished by loading them onto the column in an aqueous mobile phase, followed by a steady increase in the concentration of a strong organic modifier, like acetonitrile. Due to small differences in their structure, solute molecules will 'stick' to the column support with different strengths. At some specific mobile phase condition, a solute species will elute from the column and pass through a detector (most commonly a UV detector). Elution fractions containing a purified solute can be collected for further analysis (*e.g.* MALDI-TOF MS).

For ion exchange HPLC, the main considerations for choosing the appropriate mobile phase system are pH and salt concentration. The pH dictates the charge of the peptide or protein solute as well as the ion exchanger (stationary phase). Thus, it is important to operate at a pH where the solute and the ion exchanger have an affinity for one another (or oppositely charged). The pH should be controlled through the use of effective buffers (*e.g.* Tris-HCl in the pH range of 7.5-9.0). Solutes will 'stick' to the column with varying degrees of strength due to the differences in their charge strengths. Bound solutes can be eluted by slowly increasing the salt concentration in the mobile phase. The salt counterions (*e.g.* Na^+ or Cl^-) competitively displace solute ions from the

charged sites on the stationary phase.[78] At some specific mobile phase condition, a solute species will elute from the column and pass through a detector (most commonly a UV-Vis detector). Elution fractions containing a purified solute can be collected for further analysis (*e.g.* MALDI-TOF MS).

RP and ion exchange HPLC have been widely used to separate and analyze biomolecules. Due to its resolving power and selectivity, HPLC methods have been used to separate short peptides differing in only a few amino acids[81, 82] and variants of proteins differing only by their level of post-translational modification.[83, 84, 85] Although each mode of HPLC is powerful on its own, difficult separations sometimes require the combination of two modes, or 2-dimensional HPLC (2D-HPLC). For example, using an ion exchange step prior to RP-HPLC can significantly simplify and improve separations.[86] In this work, several HPLC techniques will be employed during the purification and analysis of the biosynthetically produced PAA products.

2.5.3 Amino Acid Compositional Analysis

Amino acid compositional analysis is used to determine the amino acid composition of a protein or polypeptide. First, the protein sample is hydrolyzed into its amino acid constituents, usually with 6 N HCl at 110°C for 24 hrs. The free amino acids are then derivatized using UV-absorbing, chromogenic or fluorescent reagents. Examples of commonly used UV-absorbing and fluorescent dyes are phenylisothiocyanate (PITC)[87] and 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate (AQC).[88] Depending on the dye used, the amino acid derivatives are then separated either by ion exchange or reversed-phase high performance liquid chromatography. The identity of a derivatized amino acid is revealed by comparing the chromatographic pattern with that of a standard.[75] The relative amount of each amino acid is determined by its relative peak area.

Some limitations of amino acid analysis include instability and high background absorbance when using some of the derivatizing dyes and amino acid instability during hydrolysis (tryptophan, cysteine, asparanine and glutamine). However, this analysis is a powerful one. For example, it can be performed accurately at the femtomole level when using fluorescent detection.[89]

2.5.4 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

The most common method of determining the molecular weight of a protein sample is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which was first introduced in 1987 by Karas and Hillenkamp.[90] This method involves mixing and co-crystallizing the protein sample with a matrix solution, such as 2,5-dihydroxybenzoic acid and nicotinic acid. The sample is ionized by irradiation with a laser beam and desorbs from the target surface. An applied voltage is applied during the laser pulse which accelerates the ions towards a time-of-flight mass analyzer. The time-of-flight analyzer measures the finite time it takes for the ions to reach the detector, and produces a full spectrum which can be used to determine the molecular weight of the protein sample. MALD-TOF MS is a robust technique that only requires femtomoles or less of a sample and is not limited by small amounts of impurities.[90]

2.6 References

[1] Fleer GJ, Cohen Stuart MA, Scheutjens JMHM, Cosgrove T and Vincent B. *Polymers at Interfaces*. Chapman & Hall, London, (1993).

[2] Fleer GJ and Scheutjens JMHM. "Block copolymer adsorption and stabilization of colloids". *Colloid Surface.* (1990); 51:281-298.

[3] Evers OA, Scheutjens JMHM and Fleer GJ. "Statistical thermodynamics of block copolymer adsorption: 1. Formulation of the model and results for the adsorbed layer structure". *Macromolecules.* (1990); 23(25): 5221-5233.

[4] Evers OA, Scheutjens JMHM and Fleer GJ. "Statistical thermodynamics of block copolymer adsorption: 2. Effect of chain composition on the adsorbed amount and layer thickness". *Faraday Trans.* (1990); 86(9): 1333-1340.

[5] de Gennes PG. "Conformations of polymers attached to an interface". *Macromolecules.* (1980); 13: 1069-1075.

[6] Marques CM and Joanny JF. "Block Copolymer Adsorption in a Nonselective Solvent". *Macromolecules.* (1989); 22: 1454-58.

[7] Wu DT, Yokoyama A and Setterquist RL. "An Experimental Study on the Effect of Adsorbing and Non-adsorbing Block Sizes on Diblock Copolymer Adsorption". *Polym. J.* (1991); 23(5): 709-714.

[8] de Laat AWM and Schoo HFM. "Novel poly(vinyl ether) block copolymers: adsorption from aqueous solutions on α-Fe_2O_3 (Hermatite) and the mechanism of colloidal stabilization". *J. Colloid Interf. Sci.* (1997); 191: 416-423.

[9] Creutz S and Jerome R. "Effectiveness of block copolymers as stabilizers for aqueous titanium dioxide dispersions of a high solids content". *Prog. Org. Coats.* (2000); 40: 21-29.

[10] Elbert DL and Hubbell JA. "Self-assembly and steric stabilization at heterogeneous, biological surfaces using adsorbing block copolymers". *Chem. Biol.* (1998); 5(3): 177-183.

[11] Tsujii K. *Surface Activity: Principles, Phenomena and Applications*. Academic Press, San Deigo, (1998).

[12] Nnanna IA and Xia J. *Protein-Based Surfactants: Synthesis, Physiochemical Properties and Applications*. Marcel Dekker, Inc., New York, (2001).

[13] Alexandridis P and Lindman B. *Amphilic Block Copolymers: Self-Assembly and Applications*. Elsevier, Amsterdam, (2000).

[14] Shao W-H, Zhang X-E, Liu H and Zhang Z-P. "Anchor-chain molecular system for orientation control in enzyme immobilization". *Bioconjugate Chem.* (2000); 11: 822-826.

[15] Carbajal ML, Smolko EE and Grasselli M. "Oriented immobilization of proteins on grafted porous polymers". *Nucl. Instr. and Meth. In Phys. Res. B.* (2003); 208: 416-423.

[16] Shi J-X, Zhang X-E, Xie W-H, Zhou Y-F, Zhang Z-P, Deng J-Y, Cass AEG, Zhang Z-L, Pang D-W and Zhang C-G. "Improvement of homogeneity of analytical biodevices by gene manipulation". *Anal. Chem.* (2004); 76: 632-638.

[17] Friedrich MG, Gieb F, Naumann R, Knoll W, Ataka K, Heberle J, Hrabakova J, Murgida DH and Hildebrandt P. "Active site structure and redox processes of cytochrome c oxidase immobilized in a novel biomimetic lipid membrane on an electrode". *Chem. Commun.* (2004); 2376-2377.

[18] Moldes C, Garcia P, Garcia JL and Prieto MA. "In vivo immobilization of fusion proteins on bioplastics by the novel tag BioF". *Appl.Environ. Microbiol.* (2004); 70(6): 3025-3212.

[19] Moldes C, Garcia JL and Garcia P. "Construction of a chimeric thermostable pyrophosphatase to facilitate its purification and immobilization by using the choline-binding tag". *Appl. Environ. Microbiol.* (2004); 70(8): 4642-4647.

[20] Kim Y-R, Paik H-J, Ober CK, Coates GW and Batt CA. "Enzymatic surface-initiated polymerization: a novel approach for the in Situ solid-phase synthesis of biocompatible polymer poly(3-hydroxybutyrate". *Biomacromolecules*. (2004); 5: 889-894.

[21] Jia Y, Talaga DS, Lau WL, Lu HSM, DeGrado WF and Hochstrasser RM. "Folding dynamics of single GCN-4 peptides by fluorescence resonant energy transfer confocal microscopy". *Chem. Phys.* (1999); 247: 69-83.

[22] Stevens MM, Allen S, Sakata JK, Davies MC, Roberts CJ, Tendler SJB, Tirrell DA and Williams PM. "pH-dependent behavior of surface-immobilized artificial leucine zipper proteins". *Langmuir*. (2004); 20: 7747-7752.

[23] Orth J, Meyer WH, Bellman C and Wegner G "Stabilization of aqueous α-Al_2O_3 suspensions with block copoymers". *Acta Polym.* (1997); 48(11): 490-501.

[24] Evans DF and Wennerstrom H. *The Colloidal Domain: Where Physics, Chemistry, Biology and Technology Meet 2nd Ed*. Wiley-VCH, New York (1999).

[25] Sigmund WM, Bell NS and Bergstrom L. "Novel powder-processing methods for advanced ceramics". *J. Am. Ceram. Soc.* (2000); 83(7): 1557-1574.

[26] Shaw DJ. *Colloid & Surface Chemistry 4th Edition*, Butterworth-Heinemann, Oxford, (2000); p220-240, Figures 8.2 and 8.11.

[27] Lewis JA. "Colloidal Processing of Ceramics". *J. Am. Ceram. Soc.* (2000); 83(10): 2341-2359.

[28] Buscall R. *Polymer Colloid*. Elsevier Applied Science Publishers, London, (1985).

[29] Napper DH. *Polymeric Stabilization of Colloidal Dispersions.* Academic Press, London, (1983).

[30] Jones DAR, Leary B and Boger DV. "The rheology of a sterically stabilized suspension a high concentration". *J. Colloid. Interf. Sci.* (1992); 150(1): 84-96.

[31] Rajaiah J, Ruckenstein E, Andrews GF, Forster EO and Gupta RK. "Rheology of sterically stabilized ceramic suspensions". *Ind. Eng. Chem. Res.* (1994); 33: 2336-2340.

[32] Napper DH. *Polymeric Stabilization of Colloidal Dispersions.* Academic Press, London, (1983); p14, Figure1.5

[33] Piirma I. *Polymeric Surfactants*. Marcel Dekker, Inc., New York. (2001).

[34] Stevens MP *Polymer Chemistry: An Introduction 3rd Ed.* Oxford University Press, New York, (1999).

[35] Husseman M, Malmstrom EE, McNamara M, Mate M, Mecerreyes D, Benoit DG, Hedrick JL, Mansky P, Huang E, Russell TP, Hawker CJ. "Controlled synthesis of polymer brushes by "Living" free radical polymerization techniques". *Macromolecules.* (1999); 32(5): 1424-1431.

[36] Gross RA and Kalra B. "Biodegradabe polymers for the environment". *Science.* (2002); 297: 803-807.

[37] Domb AJ "Biodegradable polymers derived from amino acids". *Biomaterials.* (1990); 11: 686-689.

[38] Deming TJ "Polypeptide materials: New synthetic methods and applications". *Adv. Mater.* (1997); 9(4): 299-311.

[39] Barron AE and Zuckermann RN. "Bioinspired polymeric materials: In-between proteins and plastics". *Curr. Opin. Chem. Biol.* (1990); 3: 681-687.

[40] Deming TJ "Facile synthesis of block copolypeptides of defined architecture". *Nature.* (1997); 390: 386-389.

[41] McGrath KP, Tirrell DA, Kawai M, Mason TL and Fournier MJ. "Chemical and biosynthetic approaches to the production of novel polypeptide materials". *Biotechnol. Progr.* (1990); 6: 188-192.

[42] Cappello J, Crissman J, Dorman M, Mikolajczak M, Textor G, Marquet M and Ferrari F. "Genetic engineering of structural protein polymers". *Biotechnol. Progr.* (1990); 6: 198-202.

[43] Nagarsekar A, Crissman J, Crissman M, Ferrari F, Cappello J and Ghandehari H. "Genetic synthesis and characterization of pH- and temperature-sensitive silk-elastinlike protein block copolymers". *J. Biomed. Mater. Res.* (2002); 62(2): 195-203.

[44] Padgett KA and Sorge JA. "Creating seamless junctions independent of restriction sites in PCR cloning". *Gene.* (1996); 168:31-35.

[45] McMillan RA, Lee TAT and Conticello VP. "Rapid assembly of synthetic genes encoding protein polymers". *Macromolecules.* (1999); 32: 3643-3648.

[46] Goeden-Wood NL, Conticello VP, Miller SJ and Keasling JD. "Improved assembly of multimeric genes for the biosynthetic production of protein polymers". *Biomacromolecules*. (2002); 3: 874-879.

[47] Meyer DE and Chilkoti A. "Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system". *Biomacromolecules*. (2002); 3: 357-367.

[48] Won JI and Barron AE. "A new cloning method for the preparation of long repetitive polypeptides without a sequence requirement". *Macromolecules*. (2002); 35: 8281-8287.

[49] Shuler ML and Kargi F. *Bioprocess Engineering: Basic Concepts* Prentice Hall, Englewood Cliffs, (1992); p398.

[50] Zhang SP, Zubay G and Goldman E. "Low-usage codons in *Escherichia coli*, yeat, fruit fly and primates". *Gene.* (1991); 105(1): 61-72.

[51] Makrides SC. "Strategies for achieving high-level expression of genes in *Escherichia coli*". *Microbiol. Rev.* (1996); 60(3): 512.

[52] Bingle WH, Nomellini JF, and Smit J. "Secretion of the *Caulobacter crescentus* S-layer protein: Further localization of the C-terminal secretion signal and its use for the secretion of recombinant proteins". *J. Bacteriol.* (2000); 182:3298-3301.

[53] Bornhorst JA and Falke JJ. "Purification of proteins using polyhistidine affinity tags". *Method Enzymol.* (2000); 326: 245-254.

[54] Hochuli E, Dobeli H and Schacher A. "New metal chelate adsorbent selective for proteins and peptides containing neighboring histidine residues". *J. Chromatogr.* (1987); 411: 177-184.

[55] Lavallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF and McCoy JM. "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *Escherichia coli* cytoplasm". *Bio-Technol.* (1993); 11(2): 187-193.

[56] Guan C, Li P, Riggs PD and Inouye H. "Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein". *Gene.* (1988); 67: 21-30.

[57] Henning L and Schafer E. "Protein purification with C-terminal fusion of maltose binding protein". *Protein Expres. Purif.* (1998); 14(3): 367-370.

[58] Chong S, Mersha FB, Comb DG, Scott ME, Landry D, Vence LM, Perler FB, Benner J, Kucera RB, Hirvonen CA, Pelletier JJ, Paulus H and Xu MQ. "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element". *Gene.* (1997); 192: 271-281.

[59] Smith DB and Johnson KS. "Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione". *Gene.* (1988); 67: 31-40.

[60] Holmgren A. "Thioredoxin". *Annu. Rev. Biochem.* (1985); 54: 237-271.

[61] Lu Z, Diblasio-Smith EA, Grant KL, Warne NW, LaVallie ER, Collins-Racie LA, Follettie MT, Williamson MJ and McCoy JM. "Histidine patch thioredoxins". *J. Biol. Chem.* (1996); 271(9): 5059-5065.

[62] LaVallie ER, Lu ZJ, DiBlasio-Smith EA, Collins-Racie LA and McCoy JM. "Thioredoxin as a fusion partner for the production of soluble recombinant proteins in *Escherichia coli*". *Method Enzymol.* (2000); 326: 322-340.

[63] Stoll VS, Manohar AV, Gillon W, MacFarlane ELA, Hynes RC and Pai EF. "A thioredoxin fusion protein of VanH, a D-lactate dehydrogenase from *Enterococcus faecium*: cloning, expression, purification, kinetic analysis and characterization". *Protein Sci.* (1998); 7: 1147-1155.

[64] Shi Q, Kim SY, Blass JP and Cooper AJL. "Expression in *Escherichia coli* and purification of hexahistidine-tagged human tissue transglutaminase". *Protein Expres. Purif.* (2002); 24: 366-373.

[65] Hosfield T and Lu Q "Influence of the amino acid residue downstream of $(Asp)_4Lys$ on enterokinase cleavage of a fusion protein". *Anal. Biochem.* (1999); 269: 10-16.

[66] Lavallie ER, Rehemtulla A, Racie LA, DiBlasio EA, Ferenz C, Grant KL, Light A and McCoy JM "Cloning and functional expression of cDNA-encoding the catalytic subunit of bovine enterokinase". *J.Biol.Chem.* (1993); 268(31): 23311-23317.

[67] Smith BJ "Chemical cleavage of Proteins at Methionyl Residues". *The Protein Protocol Handbook.* Humana Press, Inc. Totowa, NJ (1996).

[68] Landon, M. "Cleavage at aspartyl-prolyl bonds". *Method Enzymol.* (1977); 47: 145-149.

[69] Bornstein P and Balian G. "Cleavage at Asn-Gly bonds". *Method Enzymol.* (1977); 47: 132-145.

[70] Antorini M, Breme, U, Caccia P, Grassi C, Lebrun S, Orsini G, Taylor G, Valsasina B, Marengo E, Todeschini R, Andersson C, Gellerfors S and Gustafsson, JG. "Hydroxylamine-induced cleavage of the asparaginyl-glycine motif in the production of recombinant proteins: The case of insulin-like growth factor I". *Protein Expres. Purif.* (1997); 11: 135-147.

[71] Park HB, Pyo SH, Hong SS, and Kim JH. "Optimization of the hydroxylamine cleavage of an expressed fusion protein to produce a recombinant antimicrobial peptide". *Biotechnol. Let.* (2001); 23: 637-641.

[72] Canova-Davis E, Eng M, Mukku V, Reifsnyder DH, Olson CV, and Ling VT. "Chemical heterogeneity as a result of hydroxylamine cleavage of a fusion protein of human insulin-like growth factor I". *Biochem. J.* (1992); 278: 207-213.

[73] Hames BD and Rickwood D. *Gel Electrophoresis of Proteins*. IRL Press, Oxford, (1990).

[74] Reynolds JA and Tanford C. *Proc. Natl. Acad. Sci.* (1970); *66*: 1002-1007.

[75] Stryer L. *Biochemistry 4th Ed.* W.H. Freeman and Company, New York, (1995).

[76] Tal M, Silberstein A, and Nusser E. "Why does Coomassie Brilliant Blue –R interact differently with different proteins – a partial answer". *J. Biol. Chem.* (1985); 260(18): 9976-9980.

[77] Switzer RC, Merril CR and Shifrin S. *Anal. Biochem.* (1979); 98: 231.

[78] Cunico RL, Gooding KM and Wehr T. *Basic HPLC and CE of Biomolecules*. Bay Bioanalytical Laboratory Inc. Richmond, (1998).

[79] Aguilar MI. *HPLC of Peptides and Proteins: Methods and Protocols. Methods in Molecular Biology. Vol. 251*. Humana Press, Totowa, (2004).

[80] Lough WJ and Wainer IW. *High Performance Liquid Chromatography: Fundamental Principles and Practise*. Blackie Academic & Professional, London (1996).

[81] Guo D, Mant CT, Taneja AK, Parker JMR and Hodges RS. *J. Chromatogr*. (1986); 359: 499.

[82] Mant CT, Burke TW, Black JA and Hodges RS. "Effect of peptide chin length on peptide retention behavior in reverse-phase chromatography". *J. Chromatogr*. (1988); 458: 193-205.

[83] Kowell M, Vater J, Kluge B, Stein T, Franke P and Ziessow D. "Separation and characterization of surfactant isoforms produced by *Bacillus subtilus*OKB 105". *J Colloid Interf. Sci*. (1998); 204(1): 1-8.

[84] Karlsson G, Gellerfors P, Persson A, Noren B, Edlund PO, Sandberg C and Birnbaum S. "Separation of oxidized and deamidated human growth hormone variants by isocratic reversed-phase high performance liquid chromatography". *J. Chromatogr A*. (1999); 855(1): 147-155.

[85] Lindsay M, Gil GC, Cadiz A, Velander WH, Zhang C and Van Cott KE. "Purification of recombinant DNA-derived factor IX produced in transgenic pig milk and fractionation of active and inactive subpopulations". *J Chromatogr. A*. (2004); 1026(1-2): 149-157.

[86] Wagner, K, Racaityte K, Unger KK, Miliotis T, Edholm LE, Bischoff R and Marko-Varga G. "Protein mapping by two-dimensional high performace liquid chromatography". *J. Chromatogr A*. (2000); 893(2): 293-305.

[87] Cohen SA and Strydom DJ. "Amino acid analysis utilizing phenylisothiocyantate derivatives". *Anal. Biochem.* (1988); 174: 1-16.

[88] Cohen SA and Michaud DP. "Synthesis of a fluorescent derivatizing reagent, 6-aminoquinolyl-N-hydroxysuccinimidyk carbamate, and its application for the analysis of hydrolysate amino acids via high performance liquid chromatography". *Anal. Biochem.* (1993); 211: 279-287.

[89] Cooper C, Packer N and Williams K. *Amino Acid Analysis Protocols*. Humana Press, Totowa, (2001).

[90] Larsen BS and McEwen CN. *Mass Spectrometry of Biological Materials, 2nd Ed.* Marcel Dekker, Inc., New York, (1998).