Molecular Marine Biology and Biotechnology (1995) 4(4), 313–322

Wresting the muscle from mussel beards: research and applications

Leszek M. Rzepecki*
College of Marine Studies
University of Delaware
Lewes, Delaware 19958 U.S.A.

J. Herbert Waite
College of Marine Studies and
Department of Chemistry and Biochemistry
University of Delaware
Newark, Delaware 19716 U.S.A.

Abstract

Marine and zebra mussels secrete byssal beards to attach themselves opportunistically to hard surfaces in their environment. By doing this, they naturally earn a reputation as fouling pests. The protein precursors of byssus in mussels are being investigated in the hope not only of discovering specific measures against these marine foulers, but also to gain some insights into the technically challenging task of engineering adhesive bonds underwater. Although byssal proteins are all part of the bearded glue that bonds them to a surface, they can be subdivided into three types depending on the function that they serve in byssal threads: (1) fibrous proteins form the load-bearing cables in the core of the threads, (2) cuticular proteins form a protective coat around the cables, and (3) adhesive proteins connect the cables to a foreign surface. A flaw in any one of these will undermine a mussel's ability to attach. The fibrous proteins can be collagenous, silk-like, elastic, or any combination of these. Covering these are the cuticular proteins, which are distinguished by their surface coupling properties, tandemly repeated primary sequence, and their high content of lysine and the exotic amino acid 3,4-dihydroxyphenyl-L-alanine (DOPA). The adhesive proteins are of low molecular weight, contain DOPA, and assemble to form microcellular solids (foams). Several of these proteins are already attracting biotechnological attention as cell and tissue attachment factors, anticorrosives, and metal-sequestering reagents.

*Correspondence should be sent to this author.

Introduction

Biological organisms are a treasure trove of elegant and ingenious solutions to microengineering problems that can beset the fabrication and preservation of both natural and man-made constructs. The execution of many projects, such as glueing dissimilar surfaces underwater, might conceivably benefit from an examination of comparable problem-solving in nature (Table 1). The discovery of natural bioengineering strategies presents a challenge to the human biomimeticist[1] more complex than deducing a detailed recipe after dissecting a baked cake. However desirable the recipe might seem, it is written in the arcane languages of biochemistry and molecular biology, and thus requires considerable expense and effort to unravel. This task promises both intellectual and practical rewards for the persistent. In this essay we will concentrate on a few recent advances in the study of marine and freshwater bivalve glues and lacquers, discuss possible implications for biotechnology, and gaze into a more distant research future.

What Constitutes the "Muscle" in Mussel Beards?

Marine and freshwater mussels, and other bivalve families with similar ecologic niches, share an example of a natural adhesive stratagem with a bonding strength, versatility, and durability that biomimeticists might profitably emulate. Mussels are primarily sessile as adults, and attach themselves opportunistically to diverse hard substrata by a

[1] *Biomimeticist:* one who analyzes the biochemical foundation of biological structures, compounds, and synthetic processes with the intent to implement insights obtained in biotechnological applications. This new subdiscipline has recently undergone the usual rite of passage by spawning its own journal.

Table 1. Polyphenolic proteins in nature.

Organism	System	Bioengineering technology
Molluscs		
Mussels*	Byssus, periostracum	Adhesive, lacquer, structural foam
Platyhelminths		
Liver fluke†	Eggcase	Microencapsulating agent
Blood fluke‡		
Annelids		
Reef-building worm§	Sand tube	Cement
Tunicates		
Sea squirt‖	Blood cell	Immune system (?) Tunic repair (?)

*Waite, 1983; Waite et al., 1989; Pardo et al., 1990; Rzepecki et al., 1991; Rzepecki and Waite, 1993a, 1993b
†Rice-Ficht et al., 1992
‡Bobek et al., 1988
§Waite et al., 1992
‖Dorsett et al., 1987

complex organic structure called a byssus, colloquially known as the mussel's beard (Tamarin and Keller, 1972; Tamarin et al., 1976; Bairati and Vitellaro-Zuccarello, 1976). The beard is a bundle of up to several hundred threads that are tipped at the attachment end by disk-shaped adhesive plaques. The thread contains protein fibers in an amorphous matrix and constitutes a load-bearing cable extending from the plaque on the foreign surface to a root in the body of the mussel at the base of its beard-making organ, the foot. In the blue mussel, the byssal threads have a further complexity that is absent (or at least less obvious) in other byssate bivalves, as they can be divided into two zones of rather different mechanical properties. The smooth distal zone connected to the plaque is relatively stiff and inelastic, while the proximal zone closer to the mussel is corrugated and can be elastically stretched to more than twice its resting length. The protein fibers in the two zones are dissimilar in morphology, being coiled at the core of the proximal zone, and straight in the distal zone and the periphery of the proximal zone (Qin and Waite, 1995). This blending of mechanical properties combines strength and shock resistance in a single thread (Smeathers and Vincent, 1979). The byssal threads of bivalves such as the freshwater zebra mussels are smoother in appearance and apparently uniformly elastic, though with lesser extensibility (Eckroat et al., 1993; Rzepecki and Waite, 1993a, 1993b). These differences are consistent with the independent appearance of adult byssal attachment in several bivalve families (Morton, 1993).

At the plaque, the thread fibers are embedded in a protein resin that glues the thread firmly to the substratum. In the blue mussel, the plaque matrix has the appearance of a solid microcellular foam (Tamarin et al., 1976; Waite, 1986), which is a structure that combines mechanical strength and fracture resistance with economy of material (Rzepecki and Waite, 1991). This natural glue is remarkable for its ability to vie effectively with water for surface binding sites and thus compete away the bound water layers that interfere with adhesion (Kendall, 1994), and for its thermosetting properties, which result (in part at least) from covalent cross-linking of glue proteins following enzymatic oxidation (Rzepecki and Waite, 1991). Both features are a consequence of the hydroxylation of certain tyrosine residues in polyphenolic proteins to 3,4-dihydroxyphenyl-L-alanine (DOPA). There are at least four families of polyphenolic protein in the blue mussel plaque, but how they cooperate in the bonding enterprise is currently unclear. Several specific roles might be anticipated: foaming agents and foam stabilizers (perhaps combined in a single cross-linkable protein), priming agents at the interface between the plaque and the substratum, and bonding agents to mediate the interaction between the thread fibers and the plaque foam. The contents of the voids between the foam trabeculae are still unknown.

The thread and the plaque are rapidly assembled in a process resembling polymer injection molding (Waite, 1992). This entire microengineering process, complete in two to five minutes, requires

careful timing and precision molding, and is emphatically not a simple result of spontaneous assembly of protein precursors. Byssal precursor proteins are secreted by holocrine glands in the mussel foot and conveyed by ciliated ducts to a cylindrical mold formed by temporarily sealing the edges of a long groove in the ventral surface of the foot. The distal end of this groove splays out to form an oval cavity for plaque assembly. Following secretion, the thread fibers are finished by peristaltic action of the foot muscles, and both thread and plaque are sized by a tough lacquer. This lacquered cuticle must stick tightly to the underlying fibers and yet present a relatively inert outer surface to inhibit chemical or bacterial degradation. The cuticle is often enriched in metals (including transuranic metals), particularly in heavily polluted areas where mussel beards are used to monitor metal pollution (for a brief review, see Rzepecki and Waite, 1991). Although the metal distribution coincides with that of the polyphenolic lacquer, presumably as a consequence of the strong affinity of DOPA for multivalent metal ions (Martell, 1982), the concentration of cuticular metals seems to result from biological detoxification mechanisms rather than passive adsorption from seawater.

Fresh Mussel Beards are Chemically Active

Although the mussel beard is a nonliving extracorporeal structure, it would be misleading to picture it as chemically inert, at least in the initial stages after assembly. The integrity of both byssal lacquer and plaque foam is preserved by the action of polyphenoloxidases, which catalyze the O_2-dependent oxidation of peptidyl DOPA to dopaquinone, a process known as tanning, curing, or sclerotization (Rzepecki and Waite, 1991). DOPA as the free amino acid is familiar in the early treatment of Parkinson's disease, but its utility as an integral protein residue in mussel adhesion arises from its hydrogen-bonding and metal-complexation capacities, together with versatile covalent cross-linking potential following dopaquinone formation. Thus strong adhesive and cohesive properties are gained by the implementation of a single type of protein modification, although recent work has shown that certain polyphenolic proteins may bear further modifications. The glue and lacquer are thus two-part resins, with an emulsion of polyphenolic protein, propolyphenoloxidase (the inactive curing agent), and presumably a proteolytic propolyphenoloxidase activator, carefully packaged in the same secretory vesicles (Tamarin and Keller, 1972). This co-packaging ensures efficient mixing of the resin components and so minimizes development of potential weak points resulting from pockets of inadequate cross-linking owing to decreasing diffusivity of individual polyphenoloxidases in an increasinglycross-linked protein matrix. Consequently, it can be argued that the design of polyphenolic proteins and their polyphenoloxidases may be dovetailed to allow co-secretion in stoichiometric quantities with an assurance of effective combination in the resin. After secretion of the resin, the polyphenoloxidase is activated and the oxidation of DOPA begins. Curing can be followed microscopically in the zebra mussel byssus, which develops a yellow autofluorescence (upon excitation by blue light) during the first few hours after assembly (Rzepecki and Waite, 1993a). The chemical identity of the yellow fluorophore is unknown, but it may be derived from reaction between neighboring dopaquinone residues. The reactivity of nucleophilic protein residues with extraneous probes changes in parallel, indicating that chemical processing continues after secretion. The loss of residues to processing, however, is probably relatively minor, as noone has succeeded in measuring consistent and convincing changes in the amino acid composition of byssus with age, or detected major amounts of cross-linked products (Holl et al., 1993). Nevertheless, the sclerotized protein matrix is insoluble unless it is hydrolyzed under extreme conditions. The essentials of this cross-linking chemistry, namely, polyphenolic proteins and polyphenoloxidases, are also found in other invertebrate species (Table 1), although the details of protein structure vary widely.

Thread Fibers Contain Natural Copolymers

The characterization of the protein components of fibers in the blue mussel byssus is in progress (Qin and Waite, 1995). Two protein fragments, Col-D (60 kD) and Col-P (50 kD), have been isolated from byssus following proteolysis with pepsin, and their precursors (97 and 95 kD, respectively) detected immunologically in extracts from mussel feet. Although complete protein sequences are not yet available, current data confirm old suspicions that proteins with collagenous domains are significant structural elements (Melnick, 1958; Pujol et al., 1976; Mascolo and Waite, 1986). Col-D and Col-P have been partially characterized and occur in complementary concentration gradients along the length of the thread. Analogous gradients exist in the mussel foot, indicating that these gradients were

preformed in the collagen gland of the foot rather than established during byssal assembly. Col-D is associated with the linear protein fibers in the stiff distal zone of the thread, and in subcuticular regions of the elastic proximal zone, while Col-P appears restricted to the coiled fibers of the inner core of the proximal zone. The collagen-like sequences of Col-D and Col-P differ from mammalian collagens in that they lack 3-*trans*-hydroxyproline and hydroxylysine but resemble other invertebrate collagens in the occurrence of vicinal glycine pairs.

The compositional differences between Col-D and Col-P and their parent precursors suggest that the pre-Col-D and pre-Col-P proteins are copolymeric in nature with structurally distinct domains. The noncollagenous domains of pre-Col-D are enriched in glycine and alanine (over 30 mol% each) and resemble silk fibroin (Rudall, 1962) in overall composition. It is tempting to speculate that the stiffness of the distal thread region results from the use of a silk/collagen copolymer, with paracrystalline regions of silken β-sheet interspersed among collagen fibrils. Silk fibers are considerably tougher than collagen (Denny, 1988). The noncollagenous domains of pre-Col-P resemble insect resilin (Seifter and Gallop, 1966) in composition, and are thus part of a copolymer with rather different mechanical properties that might account for the elasticity of the proximal region. These presumably terminal domains have not been recovered from byssal extracts and may thus be targets for cross-linking. The occurrence of discontinuous pre-Col-D fibers in the subcuticular regions of the proximal thread might be necessary as a buffer to prevent shear delamination between a highly elastic core and an almost wholly inelastic cuticle during stretching. Because the proximal cuticle has a highly corrugated appearance at its resting length, models to accommodate these features are not difficult to elaborate (Qin and Waite, 1995).

The use of collagen copolymers in byssus is presently known only in marine mussels (mytilids) such as the blue mussel and may have arisen to accommodate the turbulence of the intertidal zones where water currents rapidly change force and direction. Other epibyssate mussels such as the zebra mussel and relatives (dreissenids) experience relatively more uniform tensile and shear forces on river and lake bottoms, and do not appear at significant disadvantage due to lack of byssal collagen. Endobyssate bivalves such as penshells (pinnids) live largely buried on the sea floor and perhaps do not need as sophisticated a holdfast.

The Matrix of the Adhesive Plaque is a Complex Protein Composite

The collagenous fibers of the blue mussel thread are firmly embedded in a microcellular protein foam that constitutes the plaque matrix. Several polyphenolic proteins have been unambiguously identified as plaque components, and the sequences of two have recently been elucidated. The first to be purified and partially characterized, and a major component of the matrix, was a highly disulfide bonded, tandemly repetitive 45-kD polypeptide from the blue mussel, *Mytilus edulis,* cryptically named (for the usual reasons expedient at that time) Mefp-2 (Rzepecki et al., 1992). At least 12 variants were resolved on the basis of charge density and molecular weight. However, it was not until a homologous 52-kD protein, Mgfp-2, was cloned from the closely related *M. galloprovincialis,* using sequences from Mefp-2 to construct oligonucleotide probes, that the family membership of these kindred proteins became clear (Inoue et al., 1995). The Mgfp-2 sequence consists of two terminal domains containing runs of 4–6 aspartate and glutamate residues, and one central domain containing 11 epidermal growth factor (EGF)–like modules. The EGF-like modules can be further classified on the basis of sequence similarity into 8 amino-terminal modules of 36–37 amino acids, and 3 carboxy-terminal modules which are separated by longer peptide linkers (Figure 1). Human EGF is known to have three disulfide bridges between cysteines C1–C3, C2–C4, and C5–C6 (Appella et al., 1988). The conservation of certain residues in Mgfp-2 sequences corresponding to the highly invariable A and C loops (between cysteines C2 and C3, and C5 and C6, respectively) of hEGF and EGF-like proteins suggests that the structure of the Mgfp-2 EGF-like module may be analogous. By comparison of peptide sequences from Mefp-2 and the homologous protein sequence of Mgfp-2, the sites of tyrosine

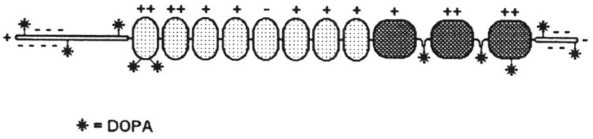

Figure 1. Schematic of the EGF-like tandemly repetitive protein Mgfp-2 from *Mytilus galloprovincialis* (Inoue et al., 1995). DOPA residues are indicated by an asterisk (*), net positive charges by a plus sign (+), and certain individual negative residues by a minus sign (−). The amino-terminus is at the left. See text for full explanation.

hydroxylation are seen to cluster at the protein termini, with 5–6 DOPA residues at each terminus. The amino- and carboxy-terminal runs of negatively charged amino acids are each sandwiched between two DOPA residues. Tyrosines are also hydroxylated in the first and last EGF-like modules, and in the peptide linkers connecting the carboxy-terminal EGF-like modules. No consensus hydroxylation sequence is obvious, and lack of hydroxylation of internal tyrosines may be a result of steric hindrance.

Extracellular and membrane proteins with EGF-like domains are common elements of the extracellular matrix and basement membranes, and are involved in cell-cell interactions and cellular differentiation (Appella et al., 1988). However, Mgfp-2 is the only known member of the EGF family that has the ability to form insoluble sclerotized matrices by virtue of the hydroxylation of tyrosine residues to DOPA. The functional specificity of EGF-like proteins is controlled by recognition sequences in the highly variable loop B (between cysteines C3 and C4). Mgfp-2 exhibits no homology in its B loop amino acid sequences with the B loops of known EGF-like proteins that are involved in cell-cell interactions, and thus seems unlikely to perform such a role. Inoue et al. (1995) showed that Mgfp-2 is expressed specifically in the foot, i.e., the organ responsible for byssal synthesis. They further showed that expression begins appropriately at the pediveliger larval stage, when larvae become competent to settle and synthesize a byssal attachment thread prior to metamorphosis. These considerations make it probable that Mgfp-2 and its homologues constitute a novel class in the EGF family, and they are currently believed to be restricted to the extraorganismic byssus.

A significant puzzle to unravel in polyphenolic protein function, especially for biomimeticists, concerns the mechanisms by which individual proteins are brought together for cross-linking. Most polyphenolic proteins are highly basic (but see below), and thus charge reduction must occur to permit protein aggregation. Laursen (1992) has proposed several potential mechanisms, including coacervate formation between the positively charged polyphenolic proteins and a hypothetical negatively charged polymer. However, pure Mefp-2 forms oligomeric aggregates during electrophoresis in sodium dodecyl sulfate (SDS) buffers that are sensitive to the degree of disulfide bond reduction, and also forms cross-linked oligomers upon prolonged storage (Rzepecki et al., 1992). Although the contribution, if any, of SDS to charge shielding in vitro is still unresolved, it is evident that the structure of Mefp-2 is sufficient for self-aggregation in the absence of another polymer. (Additional charge reduction, e.g., by salt or SDS, may be necessary for maximal condensation.) Contemplation of the sequence, probable structure, charge distribution, and hydroxylation pattern of Mgfp-2 suggests a potential mechanism for self-aggregation.

Figures 1 and 2 emphasize that the amino- and carboxy-terminal domains of Mgfp-2 have a net negative charge, while the charge density on the EGF-like domain ranges from almost neutral near its center to rather positive at its terminal modules. We can propose, therefore, that electrostatic attraction between a (negative) terminus of one Mgfp-2 molecule and a (positive) terminal module of the EGF-like domain of another molecule brings the two proteins together in an overlapping head-to-tail (or head-to-head, tail-to-tail) interaction that maximizes the cross-linking opportunity between DOPA residues in the terminal domains. This interaction appears to be specific, as it is abolished by disulfide bond cleavage. Continued accretion of polypeptide monomers would result in oligomeric aggregates, ultimately leading to an interlaced protein meshwork. If such a model proves accurate, it would be an example of a protein in which the sometimes incompatible requirements for good foamability and foam stability are successfully reconciled (Rzepecki and Waite, 1991). Because a foam is a two-phase system, the model further suggests that manipulation of the number of EGF-like repeats

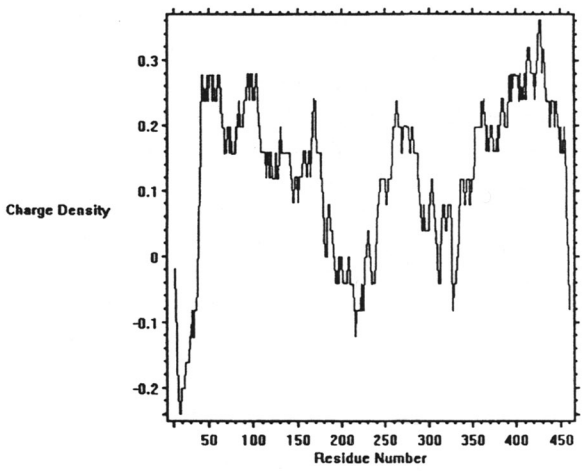

Figure 2. Charge density distribution for Mgfp-2 at pH 7.8. Charge density is given as the fractional charge per residue centered in a moving window of 25 residues.

would yield a series of polymers in which pore size of a cross-linked matrix could be controlled—a desirable feature for microencapsulation.

The other blue mussel plaque protein to have been fully characterized is the highly basic Mefp-3 (approx. 6 kD), which comes in at least nine varieties and has proved unusual in several surprising ways (Papov et al., 1995). Its most novel feature is the hydroxylation of 45% to 55% of the arginine residues to 4-hydroxyarginine, an amino acid never previously identified as a post-translational modification in proteins. Arginine/hydroxyarginine, tyrosine/DOPA, and glycine account for almost two thirds of the Mefp-3 residues (approx. 17, 22, and 25 mol%, respectively), but in contrast to other polyphenolic proteins, it is not periodic in sequence although it contains short repetitive elements. Tyrosines are almost completely hydroxylated, but no consensus hydroxylation sequence is evident. Hydroxyarginine occasionally has DOPA as a carboxy-terminal neighbor and this HOArg-DOPA bond is resistant to tryptic cleavage. As the Arg-DOPA dipeptide is trypsin labile, it seems probable that there is an interaction within the HOArg-DOPA pair, presumably a hydrogen bond. The function of Mefp-3 is unclear, but, owing to the extensive H-bonding potential of both HOArg and DOPA, it may be a surface-active agent, perhaps acting as a primer at the interface between plaque and substratum.

Byssal Lacquer: New Twists in an Old Tale

Unlike the plaque matrix, the proteinaceous lacquer that coats the external surface of thread and plaque appears to contain only one polyphenolic protein type, which, in tandem with its companion polyphenoloxidase, may constitute the bulk of the lacquer. The polyphenolic component of the blue mussel byssal lacquer is the earliest known such protein, Mefp-1 (Waite, 1983), and until recently it was believed also to constitute the glue responsible for the adhesive bond to the hard substratum (hence the older acronym MAP, "mussel adhesive protein"). Current uncertainty about the specificity of the antibodies used to localize Mefp-1 in plaques (Benedict and Waite, 1986), its evident function as a lacquer (Rzepecki et al., 1992), and the discovery of new polyphenolic plaque proteins have combined to cast doubts on a true adhesive role for Mefp-1 in nature. Nevertheless, its superlative surface-active properties (Olivieri et al., 1992), as essential for a varnish as for a glue, have evidently ensured its continued commercial availability in crude form as Cell-Tak, an adhesive preparation used in cell culture.

Mefp-1 is tandemly repetitive, consisting of many decapeptide and hexapeptide motifs containing hydroxylated amino acids, thus enjoying multiple interactions with a variety of surfaces (Waite, 1983; Filpula et al., 1990). Indeed, it has recently been shown that Mefp-1 is even more hydroxylated than first thought. The decapeptide sequence contains three proline residues, Pro-3, Pro-6, and Pro-7. Pro-7 and, occasionally, Pro-3 are hydroxylated to *trans*-4-hydroxyproline. Proline in the sixth position is occasionally hydroxylated to 4-Hyp, but was thought to be converted much more frequently to a 3-Hyp residue (Waite et al., 1985). More detailed analysis, however, has now unequivocally shown that Pro-6 is hydroxylated, not to 3-Hyp, but rather to *trans*-2,3-*cis*-3,4-dihydroxyproline, yet another novel post-translational modification (Taylor et al., 1994). Because no structural consensus for the decapeptide motif has emerged (for example, see Williams et al., 1989; Nagy et al., 1991; Trumbore, 1991; Olivieri et al., 1992; Laursen, 1992), the true significance of this correction to Mefp-1 structure is still unclear. However, this improved understanding may create more complexity for biomimeticists, as the dihydroxylation reaction may well be carried out in two steps, by two distinct prolylhydroxylases, first to 4-Hyp, then to 3,4-diHyp.

Homologous lacquer proteins from other marine mussels are curiously divergent. The lengths and sequences of their repetitive consensus oligopeptides vary widely in related species and are an evolutionary puzzle (Waite et al., 1989; Filpula et al., 1990; Laursen et al., 1990; Rzepecki et al., 1991; Inoue and Odo, 1994). Moreover, proteins from non-*Mytilus* species differ from those of *Mytilus* species in their lack of 3,4-diHyp, and have variable relative quantities of the smaller amino acids proline, glycine, serine, and aspartate/asparagine. Features common to all include multiple tandem repeats that exhibit remarkable conservatism in DNA codon usage, basicity, extensive hydroxylation, high proportion of β-turn-associated amino acids, and substantial conversion of tyrosine to DOPA; these are probably essential for function.

Byssal Proteins in Freshwater Mussels

The byssus of freshwater mussels also incorporates polyphenolic proteins. Two polyphenolic proteins have been localized to the thread and thread attach-

ment site in the plaque of the zebra mussel, *Dreissena polymorpha,* and these may constitute the lacquer or certain plaque components or both (Rzepecki and Waite, 1993a, 1993b). In keeping with the species name and appearance, one of these proteins, Dpfp-1, is known to be polymorphic with at least 8–9 variants. Neither shows any primary structural relation to mytilid or other polyphenolic proteins. They have two major claims to originality. First, although Dpfp-2 is as basic as most other known polyphenolic proteins (pI ≥ 10), Dpfp-1, unusually, is acidic (pI in the range 5.5–6.5). This is intriguing in view of the discussion above of protein aggregation mechanisms. Although neither protein has been individually localized, if they should be mapped together in the byssus, one might hypothesize the formation of a natural coaservate. Such a simple relation is complicated by the observation that the zebra mussel maintains two distinct populations of Dpfp-1, one with up to 66% tyrosine hydroxylation, the other much less hydroxylated. The population with high DOPA levels can spontaneously cross-link in solution to form oligomeric aggregates. Dpfp-2 appears to be randomly and incompletely hydroxylated, and shows no evidence of spontaneous cross-linking in similar conditions. No consensus hydroxylation sequence is evident in either protein.

The second novelty found in Dpfp-1 and Dpfp-2 is that 40% to 50% of the threonine and 10% to 30% of the serine residues are *O*-glycosylated. Most of the *O*-glycans, which have tentatively been identified as consisting of one (or at most two) *N*-acetylgalactosamines, are bound to threonine. No consensus glycosylation sequence is obvious. These are the first examples of extensively glycosylated polyphenolic proteins, and in fact it is puzzling and probably significant that other polyphenolic proteins, which are all secretory proteins, avoid glycosylation. Glycosylation may simply serve to raise the level of protein hydroxylation in this case, or it may specifically influence protein conformation. A third set of zebra mussel polyphenolic proteins contains members enriched in cystine, but does not seem glycosylated. Little is known about this group, but it may include interfacial and EGF-like proteins.

Prospects for Biotechnology

Many imaginative biotechnological applications have been proposed for the polyphenolic proteins discussed above, including, but not limited to the following:

- Medical adhesives & coatings (Nonimmunogenic; less scarring)
- Dental adhesives and fillers (more durable, versatile ceramics)
- Enzyme and cell-immobilizing agents (tissue culture, e.g., Cell-Tak; bioreactors)
- Microencapsulating agents (controlled drug, nutrient or pesticide delivery)
- Anticorrosives
- Metal scavengers (pollution monitoring and remediation; detoxification)
- Sizing agents (stronger and more durable fabrics, fibers)
- Water-resistant inks (versatile underwater surface markers).

The biotechnological attraction of exploiting polyphenolic proteins in many of these applications is rooted in an appreciation that their design has been optimized by natural selection, they have been extensively field-tested over millions of years and are highly effective, they are biodegradable, and they are nontoxic and nonpolluting. There is a current demand for improved biomaterials that are to remain in contact with tissue or body fluids (Peppas and Langer, 1994). Polyphenolic proteins are attractive from this perspective owing to their potential low immunogenicity following sclerotization. The evidence for this is partly anecdotal, based on the difficulties experienced by several investigators while trying to raise antibodies to purified polyphenolic proteins from various species. However, the sclerotized eggcases of trematodes incorporate several polyphenolic proteins (Rice-Ficht et al., 1992), and are relatively nonimmunogenic following removal of soluble surface antigens (von Lichtenberg and Raslavicius, 1967). If these proteins are such potential biotechnological marvels, it is fair to ask, now that so much is known about their structure, why more applications are not being developed.

The answer lies partly in scarce supplies of research material, partly in protein diversity and uncertainties about function, and largely in the unusual nature of the post-translational modifications necessary for correct protein function, which have thus far been found only in an eclectic subset of eukaryotic organisms. The last point is a serious stumbling-block. Cloning strategies for protein production are essential, as purification from natural sources is prohibitively expensive, but they are constrained by dismal realities. First, the large repetitive polyphenolic proteins have proved trouble-

some to clone and maintain as stable chromosomal inserts (Filpula et al., 1990; Laursen et al., 1990). A second consideration is the problem of familial protein polymorphism within a species. In helminths (parasitic flatworms), polymorphism results in large part from the existence of multigene families (Bobek et al., 1988; Rice-Ficht et al., 1992). Although allelic polymorphism has not yet been ruled out as a primary factor for known polyphenolic protein microheterogeneity in other invertebrates, it seems rather unlikely wholly to account for the extensive variation usually observed, which may be caused either by multiple genes or by differential mRNA splicing or both. It is thus conceivable that polymorphism may prove relevant to polyphenolic protein function, and perhaps it should be considered in long-term cloning strategies. Third, and most ominously, the hosts for production scale gene expression, such as *Escherichia coli* or yeast, simply cannot perform most or any of the hydroxylation and glycosylation reactions necessary, for example, to turn a merely sticky protein into a bona fide glue. Thus to mimic byssal glue (a complex of proteins) accurately could entail expression *and* purification of some or all of the tyrosyl, prolyl, and arginyl hydroxylases; a suite of polyphenolic proteins; and at least one propolyphenoloxidase and propolyphenoloxidase activator. Most of these enzymes still await discovery and characterization.

Of course, it is not the aim of a biomimeticist to reproduce a natural system verbatim, but to focus on the minimal features required to perform a particular task. There may be simplifying factors. One curious feature of some polyphenolic proteins that might ease the design of cloning strategies is the evolutionary diversity of certain protein families, exemplified by mussel lacquers, which exhibit considerable flexibility in sequence criteria. This remarkable explosion of designs implies that there may be multiple acceptable solutions for a given problem. Such versatility may even tolerate com-bination of various functionalities in chimeric proteins. Another potential amelioration of the problem might result from the apparent lack of consensus sequences to direct the peptidyl hydroxylases, glycosylases, and polyphenoloxidases. Instead, action by such enzymes may be more reflective of the solvent availability of specific residues. In many cases, an incomplete random or shotgun approach to modification may prove adequate for the task, and it is possible that the appropriate enzyme would modify a spectrum of proteins other than its natural target (residue specificity, of course, would remain unavoidable). Nevertheless, the compositional complexity of the biological resins and the essential co-packaging of resin components, coordinated secretion, and detailed post-assembly processing, together suggest that the assembly of components is as sophisticated and crucial a procedure as their initial synthesis. Biomimesis of such a system, or even a functional part of such a system, would be a formidable achievement indeed.

Some basic and applied research challenges ahead are clear, but, in the absence of immediately tangible rewards, it is increasingly more challenging to fund the necessary studies than to perform them. In an attempt to bait the hook and combine utility with intellectual curiosity, two pressing classes of problem might serve as useful foci for fundamental research with important resonances in applied research. One is the necessity to assign the correct function to each polyphenolic protein. Because functional specialization is evident, efficient biomimicry of, say, a glue would benefit from a detailed architectural map of the natural composite to identify and separate primary interfacial components from secondary structural ones. Functional assignment would allow more relevant selection of targets for applied research. The other class of problem centers on characterizing the post-translational modification mechanisms necessary for the correct function of polyphenolic proteins. This entails enumeration of the types of modification required for function (i.e., which combination of glycosylation or of tyrosine, proline, or arginine hydroxylation is essential for a particular protein), and total characterization of the requisite enzyme systems and their genes. Diverse (and easily justifiable) benefits would accrue directly from achievement of the latter goal. For example, in the case of the hydroxylases and polyphenoloxidases essential for helminth eggcase synthesis, a model system could then be designed for high-throughput screening of inhibitors that could serve as the basis for antiparasitic drug development. Such inhibitors might also aid in antifouling measures. Second, sources of modification enzymes could alleviate the scarcity of correctly modified polyphenolic proteins for research and applications by granting the ability to perform the correct modifications in vitro. More futuristically, correct protein processing might be achieved either by attempting tissue culture of the appropriate synthetic cells analogously to the production of monoclonal antibodies, or by developing methodology,

as can be done in baculovirus/insect cell cloning systems (Dodson et al., 1989), to co-express appropriate suites of proteins and enzymes. Neither is a trivial or inexpensive undertaking. These are ambitious goals, and at present seem quite distant. Still, much has been accomplished in the decade since the discovery of the first polyphenolic protein, and we are now much better situated to map the next steps.

References

Appella, E., Weber, I.T., and Blasi, F. (1988). Structure and function of epidermal growth factor-like regions in proteins. *FEBS Lett* 231:1–4.

Bairati, A., and Vitellaro-Zuccarello, L. (1976). The ultrastructure of the byssal apparatus of *Mytilus galloprovincialis*. *Cell Tissue Res* 166:219–234.

Benedict, C.V., and Waite, J.H. (1986). Location and analysis of byssal structural proteins of *Mytilus edulis*. *J Morphol* 189:171–178.

Bobek, L.A., Rekosh, D.M., and LoVerde, P.T. (1988). Small gene family encoding an eggshell (chorion) protein of the human parasite *Schistosoma mansoni*. *Mol Cell Biol* 8:3008–3016.

Denny, M.W. (1988). *Biology and the Mechanics of the Wave Swept Environment*. Princeton, NJ: Princeton University Press.

Dodson, M.S., Crute, J.J., Bruckner, R.C., and Lehman, I.R. (1989). Overexpression and assembly of the Herpes simplex virus type 1 helicase-primase in insect cells. *J Biol Chem* 264:20835–20838.

Dorsett, A., Hawkins, C.J., Grice, J.A., Lavin, M.F., Merefield, P.M., Parry, D.L., and Ross, I.L. (1987). Ferreascidin: a highly aromatic protein containing 3,4-dihydroxyphenylalanine from the blood cells of a stolidobranch ascidian. *Biochemistry* 26:8078–8082.

Eckroat, L.E., Masteller, E.C., Shaffer, J.C., and Steele, L.M. (1993). The byssus of the zebra mussel *(Dreissena polymorpha):* morphology, byssal thread formation and detachment. In: Nalepa, T.F., and Schloesser, D.W. (eds.). *Zebra Mussels: Biology, Impacts and Control*. Boca Raton, FL: CRC Press, pp. 239–263.

Filpula, D.R., Lee, S.M., Link, R.P., Strausberg, S.L., and Strausberg, R.L. (1990). Structural and functional repetitions in a marine mussel adhesive protein. *Biotechnol Prog* 6:171–177.

Holl, S.M., Hansen, D., Waite, J.H., and Schaefer, J. (1993). Solid-state NMR analysis of cross-linking in mussel protein glue. *Arch Biochem Biophys* 303:255–258.

Inoue, K., and Odo, S. (1994). The adhesive protein cDNA of *Mytilus galloprovincialis* encodes decapeptide repeats but no hexapeptide motifs. *Biol Bull* 186:349–355.

Inoue, K., Takeuchi, Y., Miki, D., and Odo, S. (1995). Mussel adhesive plaque protein gene: a novel member of the epidermal growth factor-like family. *J Biol Chem* 270: 6698–6701.

Kendall, K. (1994). Adhesion: molecules and mechanics. *Science* 263:1720–1725.

Laursen, R.A. (1992). Reflections on the structure of mussel adhesive proteins. In: Case, S.T. (ed.). *Results and Problems in Cell Differentiation. 19. Biopolymers*. Berlin: Springer-Verlag, pp. 55–74.

Laursen, R.A., Ou, J.-J., Shen, X.-T., and Connors, M.J. (1990). Characterization and structure of mussel adhesive proteins. *Materials Res Soc Symp Proc* 174:237–242.

Martell, A.E. (1982). Stability constants of metal-ion complexes. Suppl. No. 1, Part II, Special Publication No. 25, London: The Chemical Society.

Mascolo, J.M., and Waite, J.H. (1986). Protein gradients in the byssal threads of some marine bivalve molluscs. *J Exp Zool* 240:1–7.

Melnick, S.C. (1958). Occurrence of collagen in the phylum mollusca. *Nature* 181:1483.

Morton, B. (1993). The anatomy of *Dreissena polymorpha* and the evolution and success of the heteromyarian form in the Dreissenoidea. In: Nalepa, T.F., and Schloesser, D.W. (eds.) *Zebra Mussels: Biology, Impacts and Control*. Boca Raton, FL: CRC Press, pp. 185–215.

Nagy, P.I., Patel, H.C., Dreyer, W., and Hopfinger, A.J. (1991). Conformational analysis and aqueous hydration studies of model peptides for the adhesive protein of the mussel, *Mytilus edulis* L. *Int J Peptide Protein Res* 38:84–89.

Olivieri, M.P., Baier, R.E., and Loomis, R.E. (1992). Surface properties of mussel adhesive protein component films. *Biomaterials* 13:1000–1008.

Papov, V.V., Diamond, T.V., Biemann, K., and Waite, J.H. (1995). Hydroxyarginine-containing proteins in the adhesive plaques of the marine mussel *Mytilus edulis* L. *J Biol Chem* 270:20183–20192.

Pardo, J., Gutierrez, E., Saez, C., Brito, M., and Burzio, L.O. (1990). Purification of adhesive proteins from mussels. *Protein Expression Purif* 1:147–150.

Peppas, N.A., and Langer, R. (1994). New challenges in biomaterials. *Science* 263:1715–1720.

Pujol, M.M., Bocquet, J., and Borel, J.-P. (1976). Le byssus de Mytilus: étude électrophorétique des fractions protéiques riches en hydroxyproline extraites de la "glande du collagène". *C R Acad Sci* 283:555–558.

Qin, X., and Waite, J.H. (1995). Exotic collagen gradients in the byssus of the mussel *Mytilus edulis* L. *J Exp Biol* 198:633–644.

Rice-Ficht, A.C., Dusek, K.A., Kochevar, G.J., and Waite, J.H. (1992). Eggshell precursor proteins of *Fasciola hepatica*, I. Structure and expression of vitelline protein B. *Mol Biochem Parasitol* 54:129–142.

Rudall, K.M. (1962). Silk and other cocoon proteins. *Comp Biochem* 4:397–433.

Rzepecki, L.M., and Waite, J.H. (1991). DOPA proteins: versatile varnishes and adhesives from marine fauna. In: Scheuer, P.J. (ed.). *Bioorganic Marine Chemistry*, vol. 4. Berlin: Springer-Verlag, pp. 110–148.

Rzepecki, L.M., and Waite, J.H. (1993a). The byssus of the zebra mussel, *Dreissena polymorpha:* I. Morphology and in situ protein processing. *Mol Mar Biol Biotechnol* 2:255–266.

Rzepecki, L.M., and Waite, J.H. (1993b). The byssus of the zebra mussel, *Dreissena polymorpha:* II. Structure and polymorphism of byssal polyphenolic protein families. *Mol Mar Biol Biotechnol* 2:267–279.

Rzepecki, L.M., Chin, S.-S., Waite, J.H., and Lavin, M.F.

(1991). Molecular diversity of marine glues: polyphenolic proteins from five mussel species. *Mol Mar Biol Biotechnol* 1:78–88.

Rzepecki, L.M., Hansen, K.M., and Waite, J.H. (1992). Characterization of a cystine-rich polyphenolic protein family from the blue mussel, *Mytilus edulis* L. *Biol Bull* 183:123–137.

Seifter, S., and Gallop, P.M. (1966). The structure proteins. In: Neurath, H. (ed.). *The Proteins*. New York: Academic Press, pp. 155–458.

Smeathers, J.E., and Vincent, J.V.F. (1979). Mechanical properties of mussel byssus threads. *J Moll Stud* 45:219–230.

Tamarin, A., and Keller, P.J. (1972). An ultrastructural study of the byssal thread-forming system in *Mytilus*. *J Ultrastruct Res* 40:401–416.

Tamarin, A., Lewis, P., and Askey, J. (1976). The structure and formation of the byssus attachment plaque in *Mytilus*. *J Morph* 149:199–222.

Taylor, S.W., Waite, J.H., Ross, M.M., Shabanowitz, J., and Hunt, D.F. (1994). *trans*-2,3-*cis*-3,4-Dihydroxyproline, a new naturally occurring amino acid, is the sixth residue in the tandemly repeated consensus decapeptides of an adhesive protein from *Mytilus edulis*. *J Am Chem Soc* 116:10803–10804.

Trumbore, M. (1991). Investigation of the three dimensional conformation of the adhesive polyphenolic protein of *Mytilus edulis* using small angle scattering and molecular modelling. Thesis, University of Connecticut Health Center, Farmington.

von Lichtenberg, F., and Raslavicius, P. (1967). Host response to eggs of *Schistosoma mansoni*. V. Reactions to purified miracidia and eggshells and to viable and heat-killed whole eggs. *Lab Invest* 16:892–904.

Waite, J.H. (1983). Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, *Mytilus edulis* L. *J Biol Chem* 258:2911–2915.

Waite, J.H. (1986). Mussel glue from *Mytilus californianus:* a comparative study. *J Comp Physiol* 156B:491–496.

Waite, J.H. (1992). The formation of mussel byssus: anatomy of a natural manufacturing process. In: Case, S.T. (ed.). *Results and Problems in Cell Differentiation. 19. Biopolymers*. Berlin: Springer-Verlag, pp. 27–54.

Waite, J.H., Housley, T.J., and Tanzer, M.L. (1985). Peptide repeats in mussel glue proteins: theme and variations. *Biochemistry* 24:5010–5014.

Waite, J.H., Hansen, D.C., and Little, K.T. (1989). The glue protein of ribbed mussels *(Geukensia demissa):* a natural adhesive with some features of collagen. *J Comp Physiol* 159B:517–525.

Waite, J.H., Jensen, R.A., and Morse, D.E. (1992). Cement precursor proteins of the reef-building polychaete *Phragmatopoma californica* (Fewkes). *Biochemistry* 31: 5733–5738.

Williams, T., Marumo, K., Waite, J.H., and Henkens, R.W. (1989). Mussel glue has an open conformation. *Arch Biochem Biophys* 269:415–422.