



**Cite this article:** Hennebert E, Maldonado B, Ladurner P, Flammang P, Santos R. 2015 Experimental strategies for the identification and characterization of adhesive proteins in animals: a review. *Interface Focus* 5: 20140064. <http://dx.doi.org/10.1098/rsfs.2014.0064>

One contribution of 15 to a theme issue 'Biological adhesives: from biology to biomimetics'.

**Subject Areas:**

biochemistry

**Keywords:**

biological adhesion, metazoans, protein characterization

**Author for correspondence:**

Patrick Flammang  
e-mail: [patrick.flammang@umons.ac.be](mailto:patrick.flammang@umons.ac.be)

## Experimental strategies for the identification and characterization of adhesive proteins in animals: a review

Elise Hennebert<sup>1</sup>, Barbara Maldonado<sup>2</sup>, Peter Ladurner<sup>3</sup>, Patrick Flammang<sup>1</sup> and Romana Santos<sup>4</sup>

<sup>1</sup>Biology of Marine Organisms and Biomimetics, Research Institute for Biosciences, University of Mons, 23 Place du Parc, 7000 Mons, Belgium

<sup>2</sup>Molecular Biology and Genetic Engineering, GIGA-R, University of Liège, 1 Avenue de l'Hôpital, 4000 Liège, Belgium

<sup>3</sup>Institute of Zoology and Center of Molecular Bioscience Innsbruck, University of Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria

<sup>4</sup>Unidade de Investigação em Ciências Orais e Biomédicas, Faculdade de Medicina Dentária, Universidade de Lisboa, Cidade Universitária, 1649-003 Lisboa, Portugal

Adhesive secretions occur in both aquatic and terrestrial animals, in which they perform diverse functions. Biological adhesives can therefore be remarkably complex and involve a large range of components with different functions and interactions. However, being mainly protein based, biological adhesives can be characterized by classical molecular methods. This review compiles experimental strategies that were successfully used to identify, characterize and obtain the full-length sequence of adhesive proteins from nine biological models: echinoderms, barnacles, tubeworms, mussels, sticklebacks, slugs, velvet worms, spiders and ticks. A brief description and practical examples are given for a variety of tools used to study adhesive molecules at different levels from genes to secreted proteins. In most studies, proteins, extracted from secreted materials or from adhesive organs, are analysed for the presence of post-translational modifications and submitted to peptide sequencing. The peptide sequences are then used directly for a BLAST search in genomic or transcriptomic databases, or to design degenerate primers to perform RT-PCR, both allowing the recovery of the sequence of the cDNA coding for the investigated protein. These sequences can then be used for functional validation and recombinant production. In recent years, the dual proteomic and transcriptomic approach has emerged as the best way leading to the identification of novel adhesive proteins and retrieval of their complete sequences.

### 1. Introduction

Biological attachment systems can be subdivided into several groups according to the fundamental physical mechanisms underlying their operation [1]. There are many systems based entirely on mechanical principles (e.g. hooks, suckers or friction devices), whereas others rely on the chemistry of polymers and colloids (diverse types of adhesives) [2]. Biological adhesives, also comprising glues and cements, can be defined as sticky materials preventing the separation of two substrates. The surface properties of the substrates and the chemical and physical properties of the adhesive determine the strength of adhesion and cohesion [3]. Adhesives present several advantages compared to other attachment systems: (i) they are versatile, being able to bind surfaces with various chemistry and roughness, (ii) they can join dissimilar materials, and (iii) they show improved stress distribution in the joint [3]. This is why adhesive systems are ubiquitous in nature, being found in bacteria, fungi, protists, plants and animals. Although the diversity of adhesion mechanisms in animals is huge, currently

little is known of this complex phenomenon. However, in recent years, an increasing number of reports have been published on the characterization of biological adhesives [4–6].

In this review, we will focus on adhesives from animals, playing a key role in diverse functions and therefore varying widely in terms of structure, composition and capability. Adhesive systems occur in both aquatic and terrestrial animals. It is clear that the physico-chemical conditions are different underwater than in air, and therefore different kinds of adhesives will be required. However, all animal adhesive secretions are predominantly made up of proteins, although other components can be involved [4,5]. Recently, the development of new techniques allowing high-throughput characterization of protein-based materials, such as transcriptomics and proteomics, has accelerated the discovery of new adhesive proteins [7]. The aim of this review is to compile a list of the molecular methods that have been used to identify and characterize adhesive proteins in multicellular eukaryotic organisms (i.e. metazoans).

## 2. Animal protein-based adhesives

### 2.1. Diversity

From the biologists' point of view, adhesives may serve a variety of functions in metazoans: (i) attachment of an organism to a non-living surface, including dynamic attachment during locomotion and permanent fixation, (ii) attachment of one organism to another (phoresy or parasitism, prey capture), or (iii) attachment of exogenous materials together for the building of tubes, nests or burrows [6]. The evolutionary background and the biology of species on the one hand, and environmental constraints on the other hand, both influence the specific composition and mode of operation of the adhesive secretion in a particular organism. Biological adhesives can therefore be remarkably complex and involve a large range of interactions and components with different functions [4,5].

We have selected nine biological models (adult animals only) in which the complete sequence of at least one adhesive protein is known (a list of the proteins described in this review is presented in table 1). These organisms (figure 1) encompass the whole metazoan phylogeny, cover all types of habitats (marine, freshwater and terrestrial) and use protein-based adhesives for a large variety of functions (attachment, locomotion, prey capture, building and defence).

#### 2.1.1. Echinoderms

Echinoderms are among the most familiar marine creatures, and representatives, such as the sea stars (figure 1a), have become virtually a symbol of sea life. In sea stars and sea urchins, adhesion takes place at the level of a multitude of small appendages, the tube feet (or podia), and involves the secretion of an adhesive between these specialized organs and the substratum [37]. The tube feet are the external appendages of the ambulacral system of these organisms and are used in locomotion and attachment. They consist of a proximal extensible stem and a distal disc-shaped extremity. The stem allows tube foot movements while the disc mediates adhesion to the substratum through the secretion of an adhesive material from one or two types of adhesive cells [38]. Adhesion is temporary, however, and after the tube foot has become voluntarily detached, the adhesive material remains firmly bound to the substratum as a footprint [39,40]. At present, data on the

biochemical composition of echinoderm adhesive are only available for the sea star *Asterias rubens* and the sea urchin *Paracentrotus lividus* [39,41]. The water content of the adhesive material has never been measured but, in terms of dry weight, the adhesive material is mainly made up of proteins (20.6% in sea stars and 6.4% in sea urchins), carbohydrates (8% in sea stars, 1.2% in sea urchins) and a large inorganic fraction (approx. 40% in sea stars, 45.5% in sea urchins) [39,41]. In both species, potential novel adhesive proteins have been extracted [41,42]. However, only one has been completely sequenced: the protein Sfp1 from sea stars [8]. This protein, a primary constituent of the footprints, consists of four subunits, each displaying specific domains that mediate interactions with other proteins present in the adhesive material and on the tube foot surface. Sfp1 forms a structural scaffold and appears to provide cohesion to the adhesive layer [8].

#### 2.1.2. Barnacles

Barnacles (figure 1b) are marine sessile crustaceans that attach firmly and in large numbers to a variety of underwater natural substrata such as rocks, and also to man-made substrata such as ship hulls, leading to increased fuel consumption, requiring regular cleaning, and therefore causing major economic losses. In these organisms, attachment is mediated by the release of a permanent adhesive called cement [43,44]. This cement is produced by large isolated secretory cells (the cement cells) joined together by ducts which open onto the base of the animal [45–47]. The cement is composed of approximately 90% proteins with the remainder being carbohydrates (1%), lipids (1%) and inorganic material (4%, of which 30% is calcium) [48]. More than 10 proteins have been identified in the cement (cement proteins abbreviated as cp), of which six have been purified and characterized, originally from the species *Megabalanus rosa* and later from other species [44,49,50]. Among these proteins, three (cp-19k, cp-20k and cp-68k) have a surface coupling function, two (cp-52k and cp-100k) have a bulk function, and the last one (cp-16k) is an enzyme whose possible function is the protection of the cement from microbial degradation [44,49,50].

#### 2.1.3. Tubeworms

Some marine worms (figure 1c) of the family Sabellariidae are tube dwelling and live in the intertidal zone [51]. They are commonly called honeycomb worms or sandcastle worms because they are gregarious and the tubes of all individuals are closely imbricated to form large reef-like mounds. To build the tube in which they live, they collect with their tentacles particles such as sand grains or shell fragments from the water column and sea bottom. These particles are then conveyed to the building organ, which is a crescent-shaped structure near the mouth. There, the particles are dabbed with spots of cement secreted by two types of unicellular glands (cells with homogeneous granules and cells with heterogeneous granules), and they are added to the end of the pre-existing tube by the building organ [52–54]. Cement composition has been extensively investigated in the species *Phragmatopoma californica*, and consists mostly of several different proteins, a sulfated macromolecule and large amounts of  $Mg^{2+}$  and  $Ca^{2+}$  ions [34,35,54,55]. The five major cement proteins, named *Phragmatopoma* cement proteins (abbreviated as Pc-1 to Pc-5), have highly repetitive primary structures with limited amino acid diversity [34,55,56]. Pc-1, Pc-2, Pc-4 and

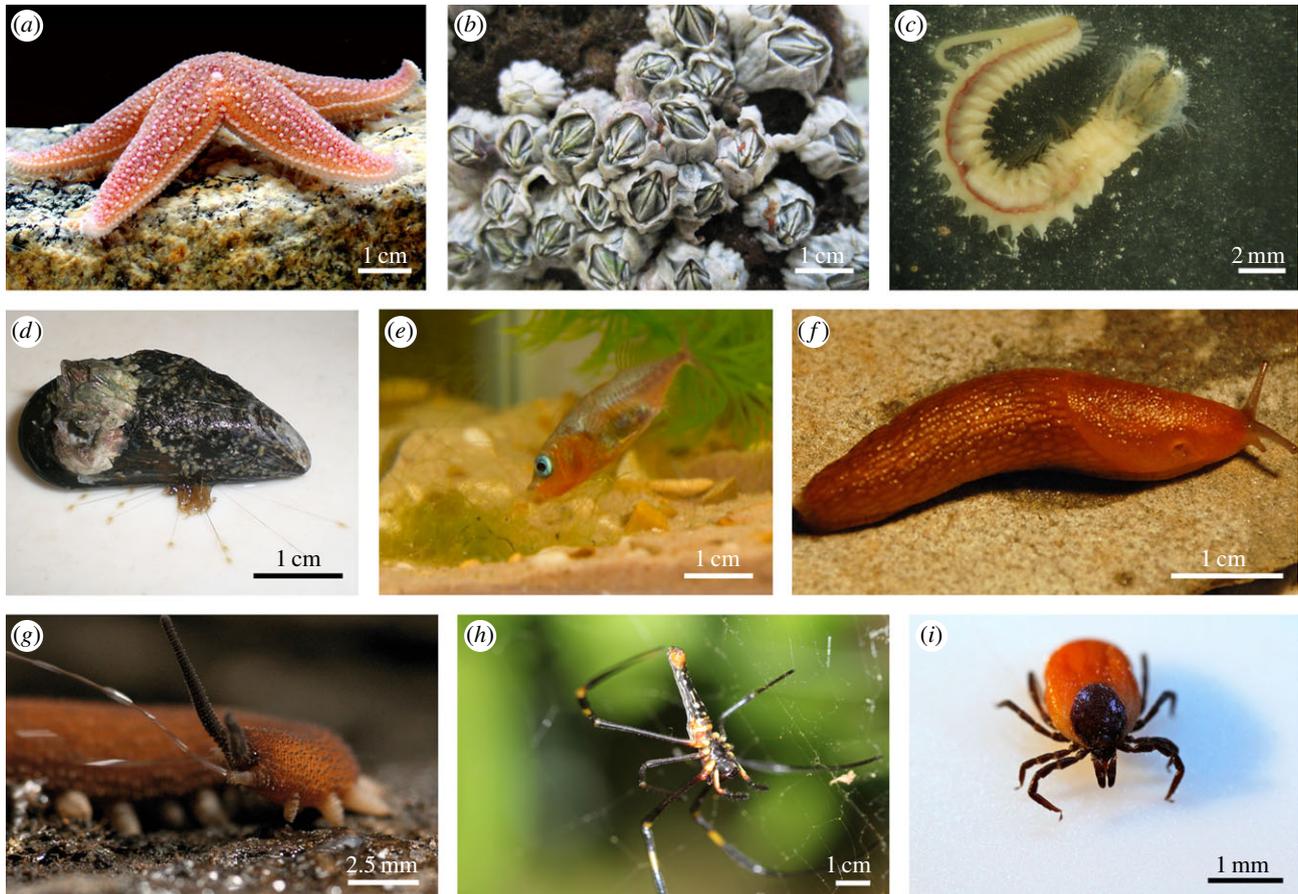
**Table 1.** List of the adhesive proteins cited in the text for which full-length sequences have been deposited in the National Center for Biotechnology Information.

common name	classification	species	protein	NCBI accession number	reference		
sea star	Echinodermata	Asteroidea	<i>Asterias rubens</i>	Sfp-1	AHN92641 [8]		
stickleback	Chordata	Actinopterygii	<i>Gasterosteus aculeatus</i>	spiggin	NM_001267690 [9]		
velvet worm	Onychophora	Udeonychophora	<i>Euperipatoides rowelli</i>	Er_P1	HM217027 [10]		
				Er_P2a	HM217028 [10]		
				Er_P2b	HM217029 [10]		
				Er_P3	HM217030 [10]		
barnacle	Arthropoda	Crustacea	<i>Megabalanus rosa</i>	Mrcp-19k	BAE94409 [11]		
				Mrcp-20k	BAB18762 [12]		
				Mrcp-52k	BAL22342 [13]		
				Mrcp-100k	BAB12269 [14]		
			<i>Balanus albicostatus</i>	Balcp-19k	AB242295 [11]		
				Balcp-20k	AB329666 [15]		
			<i>Balanus improvisus</i>	Bicp-19k	AB242296 [11]		
				spider	Araneae	<i>Nephila clavipes</i>	ASG1
			ASG2				EU780015 [16]
			PySp2				HM020705 [17]
<i>Latrodectus hesperus</i>	AgSF1	JX262195 [18]					
	AgSF2	JX262192 [18]					
	PySp1	FJ973621 [19]					
tick	Acari	<i>Rhipicephalus appendiculatus</i>	RIM36	AY045761.1 [20]			
			64P	AF469170 [21]			
		<i>Rhipicephalus haemaphysaloides</i>	RH50	AY550980 [22]			
mussel	Mollusca	Bivalvia	<i>Dreissena polymorpha</i>	Dpfp1	AAF75279 [23]		
				Dpfp2	AM229730 [24]		
			<i>Mytilus californianus</i>	Mfp-3S	DQ165556 [25]		
				Mcfp-5	DQ444853 [26]		
				Mcfp-6	DQ351537 [26]		
			<i>Mytilus edulis</i>	Mefp-1	AY845258 [27]		
				Mefp-3	AF286136 [28]		
				Mefp-5	AAL35297 [29]		
			<i>Mytilus galloprovincialis</i>	Mgfp1	D63778 [30]		
				Mgfp5	AY521220 [31]		
			<i>Perna viridis</i>	Pvfp-1	AAY46226 [32]		
				Pvfp-2	AGZ84282 [7]		
				Pvfp-3	AGZ84285 [7]		
				Pvfp-5	AGZ84279 [7]		
			slug	Gastropoda	<i>Lehmannia valentiana</i>	Sm40	ABR68007 [33]
						Sm85	ABR68008 [33]
tubeworm	Annelida	Polychaeta	<i>Phragmatopoma californica</i>	Pc-1	AAY29115 [34]		
				Pc-2	AAY29116 [34]		
				Pc-3A	AY960618 [34]		
				Pc-3B	AY960621 [34]		
				Pc-4	GH160602 [35]		
Pc-5	GH160603 [35]						

(Continued.)

Table 1. (Continued.)

common name	classification	species	protein	NCBI accession number	reference
		<i>Sabellaria alveolata</i>	Sa-1	CCD57439	[36]
			Sa-2	CCD57460	[36]
			Sa-3A	CCD57480	[36]
			Sa-3B	CCD57502	[36]



**Figure 1.** Model animals used for the study of biological adhesives. (a) Sea star of the species *Asterias rubens* attached to a rock by its tube feet. (b) Group of barnacles of the species *Elminius modestus* attached on a rock (picture courtesy of N. Aldred, Newcastle University, UK). (c) Polychaete of the species *Sabellaria alveolata* extracted from its tube. (d) Mussel of the species *Mytilus edulis* attached to a Teflon surface by means of byssal threads. (e) Three-spine stickleback of the species *Gasterosteus aculeatus* assembling its nest (picture courtesy of I. Barber, University of Leicester, UK). (f) Slug from the species *Arion fasciatus* creeping on a rock on an adhesive mucus film (picture courtesy of A. Smith, Ithaca College, USA). (g) Velvet worm of the species *Principapillatus hitoyensis* ejecting sticky threads for defence or prey capture (picture courtesy of A. Bär, University of Leipzig, Germany). (h) Spider of the species *Nephila pilipes* on its web (picture courtesy of J. Delroisse, University of Mons, Belgium). (i) Tick of the genus *Ixodes* (picture courtesy of J. Delroisse, University of Mons, Belgium). (Online version in colour.)

Pc-5 are all basic. Pc-3 is characterized by the overabundance of serine residues which are largely phosphorylated. Pc-3 is therefore extremely acidic. Tyrosine residues of both Pc-1 and Pc-2 are post-translationally modified into 3,4-dihydroxyphenylalanine (DOPA) [57]. DOPA groups take part in surface coupling either through hydrogen bonds or by forming complexes with metal ions and metal oxides present in mineral surfaces [58]. Following oxidation, DOPA groups also contribute to cement curing by forming intermolecular cross-links [34]. The different cement components are packaged and stored in concentrated granules in the two cell types. Homogeneous granules contain the sulfated macromolecules and the proteins Pc-2 and Pc-5, whereas heterogeneous granules contain the proteins Pc-1, Pc-3 and Pc-4, paired with divalent cations. Co-secretion and

limited mixing of the preassembled adhesive packets lead to formation of a complex composite cement in which the localization and role of the different adhesive proteins are still poorly understood [59].

#### 2.1.4. Mussels

Mussel is the common name used for members of several families of bivalve molluscs, from both marine and freshwater habitats. To attach themselves to the substratum, mussels produce a byssus (figure 1d), which consists of a bundle of proteinaceous threads, each connected proximally to the base of the animal's foot, within the shell, and terminating distally with a flattened plaque which mediates

adhesion to the substratum [3,27]. The composition of these plaques has been mostly characterized from marine mussels of the genera *Mytilus* and *Perna*. They are formed by the auto-assembly of secretory products originating from four distinct glands enclosed in the mussel foot. These products comprise a collagenous substance, a mucous material, a mixture of polyphenolic proteins (known as foot proteins 2 to 6, abbreviated as fp-2 to fp-6) and an accessory protein (fp-1) [3,7,26,60–62]. Since the characterization of fp-1 in the early 1980s, *Mytilus* foot proteins have been the subject of a very large number of studies leading to a detailed knowledge on their structures, functions and interactions within the byssal attachment plaque. Proteins fp-2 and fp-4 form the central core of the plaque; fp-3, fp-5 and fp-6 are located at the interface between the plaque and the substratum (primer layer); and fp-1 forms a hard cuticle protecting the core from hydrolysis, abrasion and microbial attack [27,58,62,63]. Among the latter, the presence of DOPA is a common distinctive feature shared by all the proteins identified in the byssal plaque. This modified amino acid fulfils the same roles as in tubeworm adhesive proteins: it is involved in the formation of cross-links between the different fps (cohesion) and it mediates physico-chemical interactions with the surface (adhesion) [58].

Among freshwater mussels, adhesion has been investigated mostly in zebra mussel (*Dreissena polymorpha*) because it is a highly invasive species. Although the zebra mussel byssus is superficially similar to the one of marine mussels, significant structural and compositional differences have been reported [64]. For example, contrary to the situation in marine mussels, the composition of the zebra mussel thread and plaque appears to be quite similar. To date, 13 zebra mussel proteins (Dpfp-0 to Dpfp-12) have been characterized that do display similarities to the marine fps including the presence of DOPA [23,24,65,66].

### 2.1.5. Sticklebacks

Three-spined sticklebacks, *Gasterosteus aculeatus* (figure 1e), are small freshwater fish that are used increasingly as models in evolutionary biology and ecology [67]. One fascinating aspect of stickleback biology is that the males build nests, which serve as receptacles for eggs and a focus for courtship. The male constructs his nest from plant materials, which are stuck together and to the substratum using an endogenous adhesive protein, spiggin, produced in the fish's kidney [9,68]. This protein is stored in the urinary bladder and, upon release, forms highly elastic adhesive threads. Spiggin is a cysteine-rich glycoprotein presenting a modular organization [9,68]. It is encoded by several distinct genes, some of which presenting splicing variants [69]. It is proposed that the presence of a multi-gene family coding for spiggin could contribute to the production of this protein in enough amounts to build nests during the breeding season [69]. Recently, Seear *et al.* [70] showed that spiggin genes expression is significantly affected by both the flow regime experienced by the fish and its nesting status.

### 2.1.6. Slugs

Slugs are terrestrial gastropod molluscs (figure 1f) which attach to surfaces by secreting a viscous mucus film on which they creep by mean of waves of muscular contractions running along their foot [71,72]. The so-called trail mucus is left behind by the slug

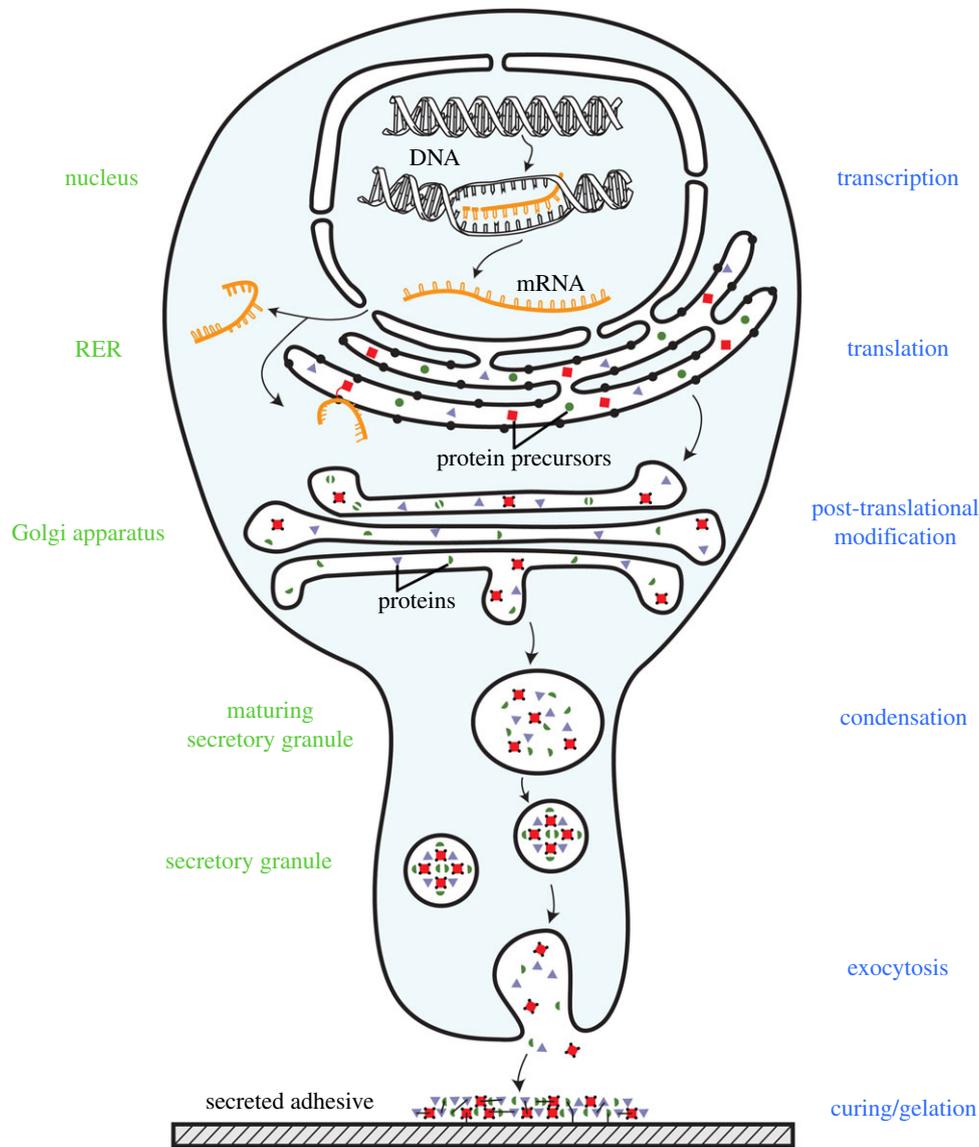
as it moves. In *Lehmannia valentiana*, this adhesive mucus is principally composed of water (approx. 90%) while the remaining 10% consists mainly of carbohydrates (34–41% of dry weight) and proteins (25–34% of dry weight) [33]. Among the 18 proteins making up the protein fraction, two have been sequenced, Sm40 and Sm85. Both contain specific domains which could promote protein–protein interactions [33].

### 2.1.7. Velvet worms

Velvet worms are terrestrial soft-bodied invertebrates inhabiting tropical and temperate forests. To defend themselves against predators or to capture their prey, these organisms use a sticky secretion produced by large glands located on each side of the gut within the body cavity and expelled via a pair of modified limbs called slime papillae (figure 1g) [73,74]. This secretion is ejected as adhesive threads forming a net entangling the prey [75,76]. To date, biochemical analyses have focused on the composition of the slime produced by velvet worms from the genus *Euperipatoides*. This slime is principally composed of water (approx. 90%) while the remaining 10% consists mainly of proteins (55% of dry weight) associated to carbohydrates (1.3%), and small quantities of lipids and the surfactant nonylphenol [75]. In the species *Euperipatoides rowelli*, the protein fraction is mainly made up of high molecular weight proline-rich proteins (Er\_P1, Er\_P2 and Er\_P3). Carbohydrate-binding proteins and small peptides are also present in the adhesive secretion where they could act as antimicrobial agents and protease inhibitors [10].

### 2.1.8. Spiders

Spiders (figure 1h) are able to spin high performance silk fibres that they use for a wide range of functions, including prey capture, locomotion and protection of eggs. In Araneomorphae, seven distinct glands comprise the spinning apparatus. Among them, two are involved in adhesion: the pyriform gland produces attachment disc silks, which attach dragline silk to substrates, and the aggregate gland produces an aqueous glue which covers the silk fibres of the web as well as an adhesive which lash the connection joints of the web [16,18,19]. The pyriform secretion from orb web and cobweb spiders is a gelatinous substance, which dries to form a strong chemical adhesive [19]. It is composed of adhesive proteins (pyriform spidroins 1 and 2) that contain different internal block repeats (including proline-rich segments) and share a high percentage of polar amino acids within their protein sequences [17,19,77]. The viscid glue produced by the aggregate glands of orb-weaving spiders functions primarily to retain flying insects. It is a complex assembly of glycoproteins and an aqueous solution of low molecular weight hygroscopic salts that regulate water content in the drop and keep the glycoproteins soft and tacky to maintain the stickiness in variable humidity environments [78,79]. The glycoprotein component is composed of two unique protein subunits (aggregate spider glue 1 and 2) that are both glycosylated. ASG1 has a high proportion of charged amino acids and is highly similar to chitin-binding proteins, while ASG2 has similarities with elastin and is thus associated with elasticity [16]. In the cobweb spider *Latrodectus hesperus*, the aggregate gland also produces two proteins, AgSF1 and AgSF2, that interface with dragline silks to form the connection joints of the web [18]. Both proteins are non-glycosylated and present internal amino acid block repeats.



**Figure 2.** Schematic illustration of the synthesis and secretory pathway followed by adhesive proteins in a typical adhesive cell (not to scale). The names of the different molecules, cellular compartments and processes involved are indicated in black, green and blue, respectively. See text for a detailed description.

### 2.1.9. Ticks

Ticks are small, ectoparasitic acarids (figure 1*i*), living by haematophagy on the blood of mammals and birds. To feed on their hosts, Ixodid ticks rely on the secretion of a cement to remain attached on the host skin for the complete duration of their blood meal. The formed cement cone extends from the tick hypostome to the host epidermis. It is primarily proteinaceous, and also contains some carbohydrate and lipid [20,80]. The protein fraction appears to be made up of proteins rich in glycine, the number of which varies according to the species [20,80–82]. Some of these proteins, generally defined as glycine-rich proteins, have been characterized and tested as anti-tick vaccines [20,80,81]. In *Rhipicephalus appendiculatus*, two cement proteins, RIM36 and 64P, have been characterized. They possess a high content of glycine, serine and proline, and contain two types of glycine-rich amino acid repeats [20,21].

### 2.2. Biosynthesis

The biosynthesis, packaging and release of the adhesive proteins by the adhesive organs follow the so-called regulated

secretory pathway [83]. This pathway is summarized in figure 2 for a typical adhesive cell. In the nucleus, the genes encoding adhesive proteins are transcribed into mRNAs, which are then matured and exported to the cytosol where they are translated at the level of ribosomes. The rough endoplasmic reticulum (RER) then captures secreted proteins from the cytosol as they are being synthesized. Indeed, these proteins possess a polypeptide signal sequence that directs the engaged ribosome to the endoplasmic reticulum membrane. The newly formed protein is then fully translocated across the RER membrane and released into its lumen. All adhesive protein precursors described so far present such a signal peptide, which is cleaved off in the RER. Proteins are then transferred from this compartment to the Golgi apparatus and from the Golgi apparatus to immature secretory granules by means of transport vesicles. As the granules mature, their contents become concentrated, probably as the result of both the continuous retrieval of membrane and the progressive acidification of the granule lumen. In adhesive cells, it has been suggested that protein condensation could involve a process called complex coacervation which is the spontaneous fluid–fluid phase separation of an aqueous protein

solution into two immiscible aqueous phases, a dilute equilibrium phase and a denser, protein-rich coacervate phase [84]. In sandcastle worms, given the presence of both polyanions (acidic proteins) and polycations (basic proteins) in the granules, coacervation is driven by the electrostatic attraction and neutralization of these oppositely charged polyelectrolytes [34,55]. In mussels, polyanions are not known to be involved in adhesion but it has been demonstrated that a zwitterionic variant of the protein fp-3 (Mfp-3S [85]) can coacervate with itself through both electrostatic and hydrophobic interactions [86]. Because their final mature secretory granules are so densely filled with contents, adhesive cells can release large amount of material promptly by exocytosis when triggered to do so.

During their transfer from one compartment to the next, proteins are successively modified. Post-translational modification (PTM) is indeed a common feature of many adhesive proteins (see §3.3). One such modification, *N*-glycosylation, takes place in the endoplasmic reticulum, with subsequent oligosaccharide processing in the Golgi apparatus [83]. Many other PTMs, such as *O*-glycosylation or phosphorylation, occur in the Golgi apparatus [83]. Another possible PTM is protein cleavage. In sea stars, cleavage of the adhesive protein Sfp-1 apparently is autocatalytic, occurring at low pH in the late secretory pathway (in the Golgi apparatus or in the maturing secretory granules) of the tube foot adhesive cells, and generating four protein subunits [8].

Upon release, the adhesive proteins spread readily on the substrate where they auto-assemble to form the adhesive joint. This formation is usually accompanied by a gelation or curing process of the adhesive. For terrestrial organisms in which the adhesive is released in the form of a liquid protein solution, gelation could be triggered by water loss owing to evaporation. This model has been proposed to explain the formation of velvet worm slime, in which water evaporation would bring hydrophobic domains and regions of opposite charge from the proline-rich adhesive proteins in closer proximity to enable ionic and hydrophobic interactions [10]. In aquatic organisms, gelation could result from a pH or ionic strength differential between the secretory granules and water. Many marine adhesive proteins are post-translationally modified with different chemical groups. Among those, phosphates and sulfates may be involved in non-covalent adhesive and/or cohesive interactions, possibly through  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$  bridging. In the cement of the tubeworm *P. californica*, which contains the polyphosphorylated protein Pc-3 [34], the jump in pH accompanying secretion, from about 5 in the secretory granule to 8.2 in seawater, could trigger a change in bonding between  $\text{Ca}^{2+}$  and phosphate from electrostatic to ionic, the effect of which would be to harden spontaneously and solidify the adhesive [87].

Creating covalent cross-links between constitutive macromolecules is another way to cure the adhesive secretion. In DOPA-containing adhesives, such as the tubeworm cement or the mussel byssal plaque, cross-linking reactions follow the oxidation of DOPA to DOPA-quinone, a reaction catalysed by a polyphenol oxidase (catechol oxidase<sup>1</sup>). Once formed, DOPA-quinone is capable of participating in a number of different reaction pathways leading to intermolecular cross-link formation [7,34,88]. Secreted catechol oxidases have been detected in both the tubeworm cement and the mussel byssus [55,89]. In barnacles, Dickinson *et al.* [90] proposed that cement polymerization could depend on glutamyl-

lysine cross-linking mediated by a transglutaminase but this hypothesis was later refuted by Kamino [50].

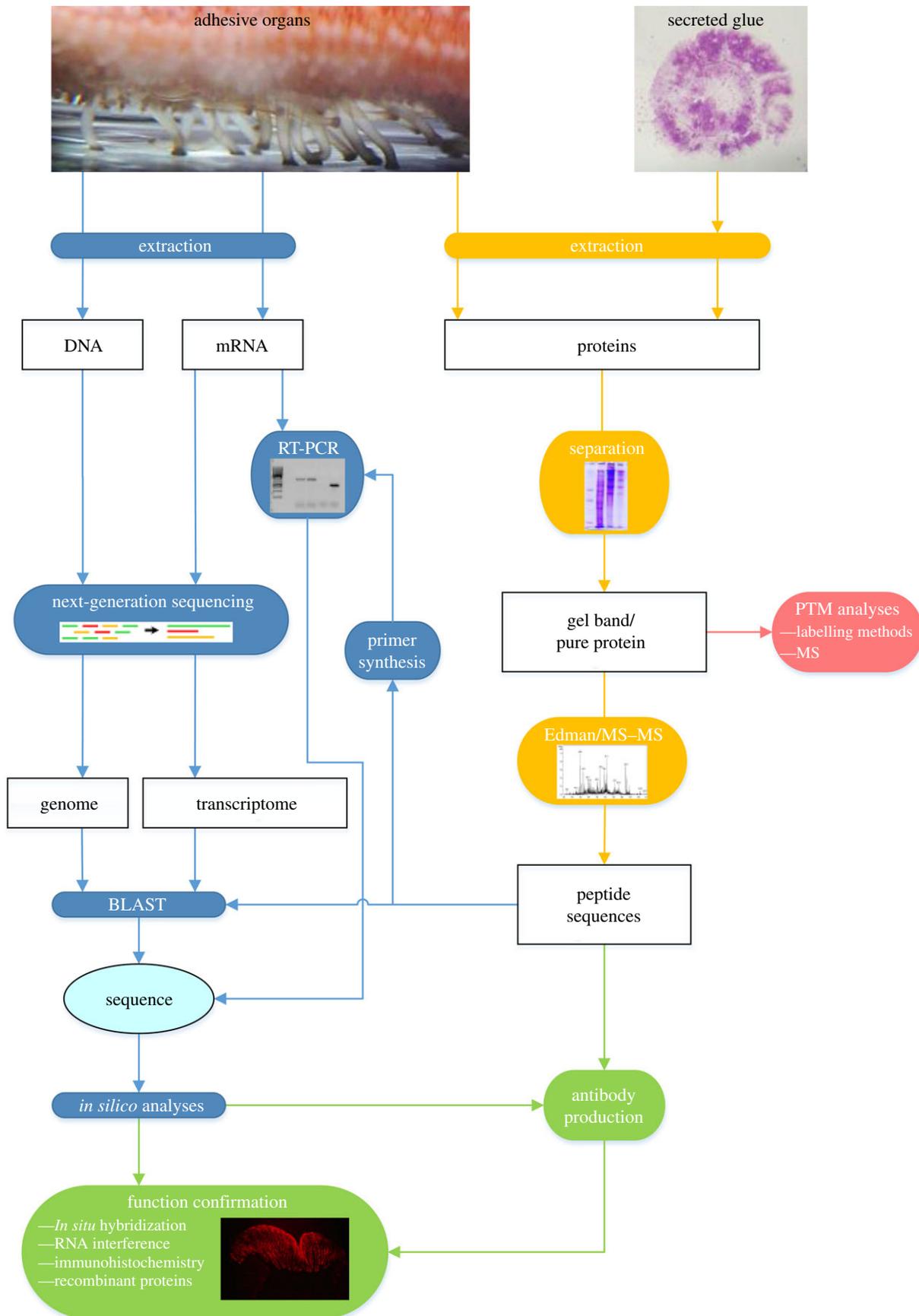
### 3. Molecular tools to characterize protein-based adhesives

On the basis of the adhesive cell secretory pathway (figure 2), two different strategies can be used to characterize putative adhesive proteins: molecular biology tools allow the retrieval of complete primary sequences from the cell nucleotidic information (DNA or mRNA) while protein chemistry tools permit one to obtain polypeptidic sequences and PTM information directly from the secreted proteins or from their precursors in the cells (figure 3). Although these two strategies can be used independently (e.g. protein chemistry approach for the mussel protein Mefp-3 [91]; molecular biology approach for the tubeworm proteins Sa-1 to Sa-3 [36]), they are generally conducted in parallel because the combination of their respective results facilitates the identification and characterization of the adhesive protein candidates.

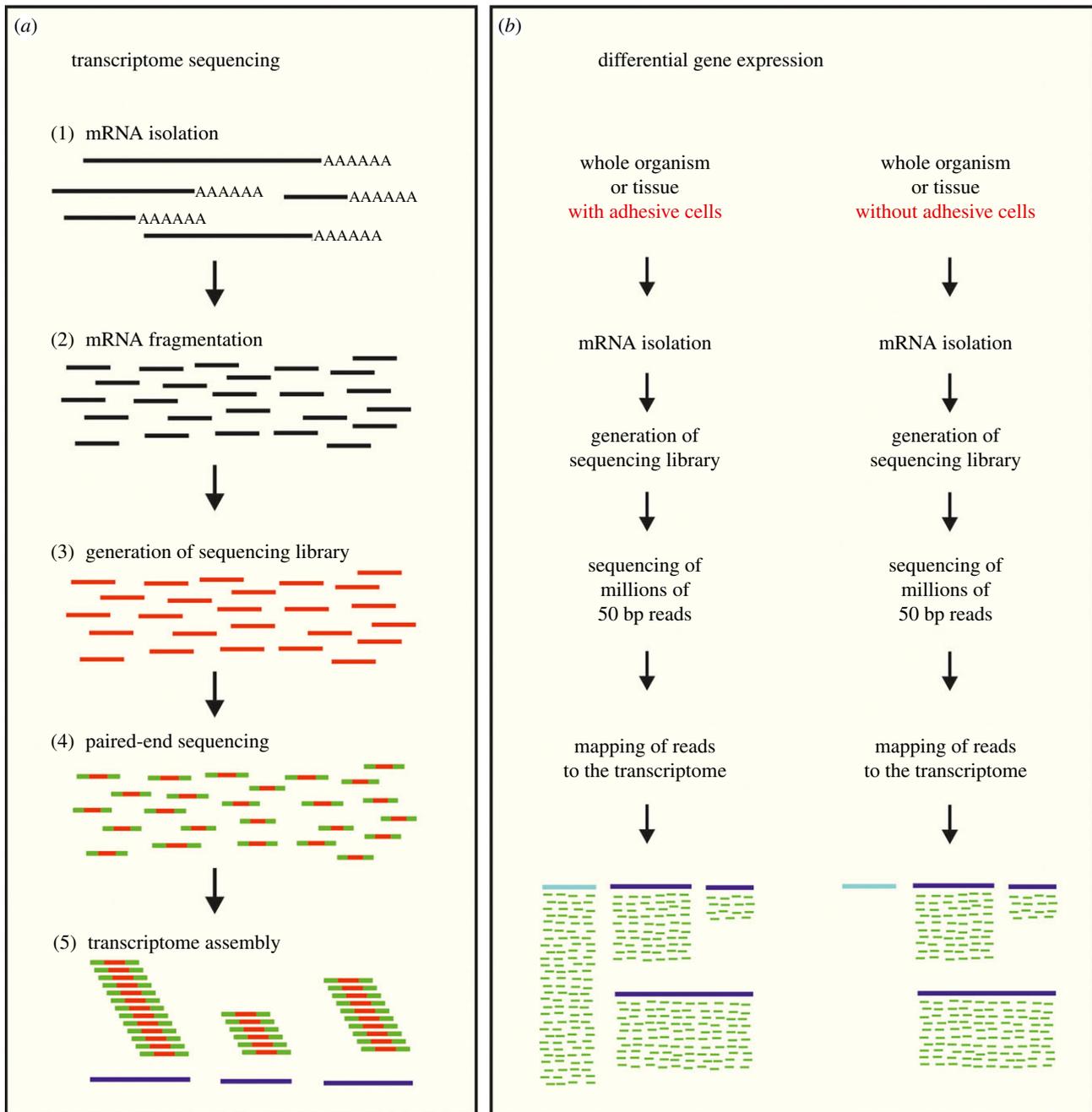
#### 3.1. Identification of adhesion-related genes and mRNAs

To the best of our knowledge, no adhesive protein has ever been identified directly from an animal's genome. The identification and isolation of adhesion-related genes can indeed be hampered by the lack of information or sequence data for the respective organism. On the other hand, for many organisms, there is no genomic database available. Therefore, mRNA sequencing is an essential prerequisite to get access to the genes expressed in certain tissue. Until recently, this was achieved by constructing a complementary DNA (cDNA) library and sequencing clones randomly to generate expressed sequence tags (ESTs; e.g. [92]). However, with the advent of next-generation sequencing (NGS), access to mRNA information is now obtained directly by transcriptome sequencing (e.g. [93,94]).

A cDNA library is a combination of cloned cDNA fragments inserted into a collection of host cells (generally bacteria), which together constitute some portion of the transcriptome of the organism or the tissue investigated. The generation of cDNA libraries involves the following steps: (i) isolation of mRNA, (ii) reverse transcription of mRNA into cDNA, and (iii) ligation of cDNA fragments into bacterial plasmids. Stewart and co-worker [35] took advantage of a cDNA library constructed from the adhesive gland of the tubeworm *P. californica* to identify potential novel adhesive proteins. As mRNAs are present in high number of copies for proteins produced continuously by the adhesive cells, they hypothesized that cDNAs coding for adhesive proteins should be more represented in the library than other cDNAs. They sequenced about 300 random clones and analysed the sequences using bioinformatics tools (see below). This approach allowed the identification of 18 genes encoding proteins with characteristics of secreted structural proteins, including the most abundant cement proteins Pc-1 to Pc-5 [35,56]. Some specific knowledge about the composition of the adhesive secretion can also be useful to obtain the full-length sequence of particular adhesive proteins. This is exemplified by the work of Zhao *et al.* [34], who compared the amino acid analysis of whole tubeworm cement to the amino acid composition of its two



**Figure 3.** Molecular tools used to characterize protein-based adhesives, illustrated by examples from sea stars. Nucleic acids (DNA and mRNA; blue pathway) are extracted from the adhesive organ(s) (here, tube feet). Both can be submitted to next-generation sequencing to obtain, respectively, the genome of the animal or the transcriptome of the adhesive organ(s). Proteins (yellow pathway) are extracted from the secreted material (here, adhesive footprints) or the adhesive organ(s). After purification, they can be analysed for the presence of post-translational modifications (PTMs; red pathway) and/or submitted to peptide sequencing by Edman degradation or tandem mass spectrometry (MS/MS). The peptide sequences can be directly used for a BLAST search in the genome or transcriptome, or to design degenerate primers to perform RT-PCR, both allowing the recovery of the sequence of the cDNA coding for the investigated protein. The sequence is then analysed *in silico* to extract information required to perform experiments allowing the validation of the adhesive function of the protein (green pathway).



**Figure 4.** Consecutive steps for transcriptome sequencing (a) and experimental set-up for differential gene expression (b). See text for details. (a) mRNA (black); sequencing library (red), note adapters at the ends are not shown; paired-end reads (green–red–green), each sequenced stretch at both ends (green) of the DNA fragment (red) is about 100 bp assembled transcripts (purple lines). (b) Transcripts of transcriptome (purple lines) and adhesive transcript (blue line). Note that no reads can be mapped to the adhesive transcript when tissue without adhesive cells used.

constitutive proteins known at that time (Pc-1 and Pc-2), and, by doing this, deduced that a third protein containing more than 80 mol% of serine was probably present in the cement. Therefore, they designed a degenerate primer corresponding to five consecutive serine residues (using the codon preference for serine in Pc-1 and Pc-2), and used it to amplify by PCR the sequence of the protein Pc-3 from a cDNA library of the tubeworm cement gland. Choresh *et al.* [16] used similar methods (random sequencing and PCR) to screen the aggregate gland cDNA library from the spider *Nephila clavipes*. They identified an abundant clone coding for two proteins: ASG1 and ASG2. The selection of clones representing highly expressed proteins can also be done through DNA microarray analysis (e.g. identification of the proteins AgSF1 and 2 from the aggregate gland of the spider *L. hesperus* [18]). In ticks, a 15 kDa cement protein,

64P, was identified in a cDNA library prepared from the salivary glands of adult female *R. appendiculatus*, based on its amino acid composition [21,95].

Generation of a transcriptome using a short reads NGS method involves several steps (figure 4a): (i) isolation of mRNA, (ii) fragmentation of mRNA, (iii) generation of a sequencing library with adapters on both ends, (iv) paired-end sequencing, and (v) bioinformatic assembly of paired-end reads. While a transcriptome provides the full complement of expressed transcripts, an adhesion-related tissue- or cell-type specific candidate gene list would be favourable. For this purpose a differential RNA-seq approach [96–99] can be applied (figure 4a), especially if only small amounts of adhesive tissue can be collected. Briefly, tissue containing adhesive cells and tissue lacking adhesive cells are sequenced using

short reads (50 bp). These reads are mapped to the existing transcriptome to identify differentially expressed transcripts, both on a qualitative and on a quantitative level [100]. This differential approach allows one to identify candidate adhesive genes in the absence of protein information (molecular weight, composition or sequence), and is therefore of particular interest when the amount and/or insolubility of the adhesive material hinders protein extraction (see also below). In the marine flatworm *Macrostomum lignano*, the differential RNA-seq approach allowed the identification of an intermediate-filament gene which, although it does not code for an adhesive protein, is essential for the adhesive process [101]. In barnacles, high-throughput NGS has been applied to compare transcriptomes of the prosoma (non-adhesive tissue) and the basis (which contains the cement gland) of the membranous-based species *Tetraclita japonica formosana* [102]. The authors showed that homologues of barnacle cement proteins cp-19K, cp-52K and cp-100K as well as adhesion-related genes such as crustin or fibroin were predominately expressed in the basis. Likewise, existing adhesive protein sequences available in publicly accessible databases can be very valuable for the identification of adhesion-related genes in related species through bioinformatic comparisons. For instance, Guerette *et al.* [7] sequenced the transcriptome of the foot of the green mussel *Perna viridis*, and conducted BLAST homology searches of all known *Mytilus* foot proteins against it. By doing so, they retrieved the full-length sequences of two adhesive proteins from *P. viridis*: Pvfp-1 and Pvfp-2. Other protein characteristics such as molecular weight or composition can also be used to help retrieve sequences from transcriptomes (e.g. Pvfp-3, Pvfp-5 and Pvfp-6 [7]).

When the sequence of an adhesive protein has already been characterized in a closely related species and high protein homology is likely, straightforward techniques of molecular biology such as RT-PCR can be used to obtain the full-length sequence of putative adhesive proteins without the need to sequence a high number of transcripts (figure 3). For example, Becker and co-workers [36] designed primers from the known cDNA sequences coding for the proteins Pc-1, Pc-2 and Pc-3A and Pc-3B of the tubeworm *P. californica* to target the cDNA coding for the homologous proteins Sa-1, Sa-2 and Sa-3A and Sa-3B in *Sabellaria alveolata*. In barnacles, a similar approach was used to identify the protein homologous to Mrcp-19k from *M. rosa* in *B. albicostatus* (Balcp-19k) and *B. improvisus* (Bicp-19k) [11]. Advantage can also be taken of the presence of repeated amino acid sequences in the known adhesive proteins. Degenerate primers can be designed on the basis of these sequences and be used in PCR in order to identify the cDNA sequence of homologous proteins in close species (e.g. the protein from *Mytilus galloprovincialis* Mgfp1 homologous to *Mytilus edulis* fp-1 [30]; and the protein from *B. albicostatus* Balcp-20k homologous to *M. rosa* cp-20k [15]). These different strategies require however that the adhesive proteins from the different species present sufficient sequence identity between each other.

Once the complete primary sequence of a putative adhesive protein has been obtained, its characterization usually starts by *in silico* analyses. A large number of bioinformatics tools are available for making predictions about the function and physico-chemical properties of a protein on the basis of its primary sequence (see, for instance, the ExPasy proteomics server [103]). These tools are of particular interest in the field of bioadhesion, where sequence

characteristics are generally recognized as major contributors to protein adhesive properties. One major tool used in most studies is SignalP [104], which predicts the presence of the signal peptide, a short (5–30 amino acid long) hydrophobic sequence present at the N-terminal part of proteins destined to be secreted (see §2.2). Identification of the so-called secretome is an approach that has been used in tubeworms, for example [35]. The analysis of amino acid composition of a protein (e.g. with ProtParam [105]) allows one to reveal the presence of highly expressed amino acids (e.g. the glycine- or serine-rich proteins in the tubeworm *P. californica* [35]; or the glycine-rich proteins of ticks [80]) and the resulting isoelectric point (pI) of the protein (e.g. the basic proteins Sa-1 and Sa-2 in the tubeworm *S. alveolata* [36]). The presence of repeated amino acid sequences is a frequent feature in adhesive proteins that can be identified using the Rapid Automatic Detection and Alignment of Repeats tool (RADAR [106]). For instance, this analysis has been used to identify the tandem repeats made up of consecutive charged residues interspersed with short hydrophobic regions in the velvet worm proteins Er\_P1 and Er\_P3 [10]. In some species, adhesive properties appear to result from the presence of conserved structural domains able to link to other molecules, and which can be identified using, for example, the Conserved Domain Database from NCBI [107]. Indeed, many adhesive proteins display specific protein-, carbohydrate- and metal-binding domains: e.g. von Willebrand factor type D domain in the stickleback protein spiggin [68,69]; EGF-like and von Willebrand factor type A domains in the slug proteins Sm40 and Sm85 [33]; discoidin, von Willebrand factor type D, galactose binding, C8 and EGF-like domains in the sea star protein Sfp1 [8]. Finally, a number of tools are also available to predict potential sites of PTMs such as glycosylation, hydroxylation and phosphorylation in the primary sequence. Examples of the use of these tools can be found in Kawahara & Nishida [9] for the fish protein spiggin, and in Haritos and co-workers [10] for velvet worm adhesive proteins.

### 3.2. Identification of adhesive proteins

To date, most studies on adhesive proteins have been carried out on proteins isolated from two enriched sources: the secreted adhesive material or the adhesive gland/organ (figure 3). Although in some cases proteins can be easily isolated from the adhesive material (e.g. in velvet worms [75]), protein extraction is generally hindered by the insolubility of the adhesive material. Strong denaturing (e.g. urea, guanidine hydrochloride) and/or reducing (e.g.  $\beta$ -mercaptoethanol, dithiothreitol) conditions were indeed required to isolate proteins from spider webs [18], stickleback nest threads [69], mussel byssal adhesive plaques [108], barnacle cement [14] and sea star and sea urchin adhesive footprints [41,42]. To bypass the challenge of solubilizing the secreted cured adhesive material, it is sometimes more convenient to perform protein extraction on the dissected adhesive gland or organ, which are a source of soluble adhesive precursors (e.g. the stickleback kidney [68]; the mussel foot [26] or the sea urchin tube feet [109]). Once in solution, adhesive proteins must be purified, i.e. separated from non-protein parts of the mixture and from all other proteins. Purification steps exploit differences in protein size (e.g. separation by Sodium dodecylsulfate polyacrylamide gel electrophoresis [33]), size and isoelectric point (e.g. separation by two-dimensional polyacrylamide gel electrophoresis

[110]), physico-chemical properties (e.g. precipitation with ammonium sulfate [108] or acetone [111]) and binding affinity (separation using different types of chromatography [57]).

Usually, the next step in adhesive protein identification is obtaining peptide sequences. Although several techniques exist to obtain partial amino acid sequence of a protein, the two dominant methods are Edman degradation and mass spectrometry (MS). Both methods have been successfully used to obtain the sequence of peptide fragments from adhesive proteins. These peptides result from the digestion of the purified protein (either in solution or in a gel band/spot) using chemical agents (e.g. cyanogen bromide [14]) or endopeptidases (e.g. trypsin or pepsin [41,112]). Peptides are usually desalted and/or separated by electrophoresis and/or chromatography before sequencing [109]. The Edman degradation method involves the chemical modification of the N-terminal amino acid of the peptide, which can then be cleaved from the chain and identified. Successive cycles of degradation are then used to identify the sequence of the peptide (e.g. mussel fps [108,111]). The same method can also be exploited to determine the N-terminal amino acid sequence of pure proteins (e.g. the tubeworm protein Pc-2 [57]) or of proteins separated by electrophoresis, transferred onto poly(vinylidene fluoride) or nitrocellulose membranes, and excised (e.g. the barnacle protein Mrcp-20k [12]).

In MS, biomolecules are ionized and their mass is measured by following their specific trajectories in a vacuum system [113]. Two fundamental strategies for protein identification and characterization by mass spectrometry are currently employed: top-down proteomics, in which intact proteins or large protein fragments are subjected to gas-phase fragmentation for MS, and a bottom-up proteomics, in which purified proteins, or complex protein mixtures, are subjected first to proteolytic cleavage, and what is analysed by MS are the peptide products (see [114–117] for review). Although MS can measure the mass of intact adhesive proteins (e.g. in mussel adhesive footprints [7,25]), it is usually used to sequence peptides resulting from enzymatically digested adhesive proteins. The two most popular techniques to identify protein sequences using MS are peptide mass fingerprinting (PMF) and tandem mass spectrometry (tandem MS or MS/MS). In the first method, the mass spectrometer generates a list of peptide masses which are compared with calculated peptide masses generated by *in silico* cleavage of protein or cDNA sequences present in publicly available databases using the same specificity as the enzyme that was employed experimentally (e.g. NCBI [118]). PMF can be performed with the same instrumentation used for MS/MS, but it is usually done using time-of-flight mass spectrometers with matrix-assisted laser desorption ionization. This method is however seldom used to identify adhesive proteins, because it requires pure proteins or simple mixtures of proteins, and several peptides are needed to unambiguously identify a protein. Tandem MS allows *de novo* internal sequencing, in which isolated peptides are fragmented and the generated mass spectra of resulting fragments are analysed to reconstitute the peptide sequences (e.g. slug proteins [33]; sea urchin and sea star proteins [41,42,109]; zebra mussel proteins [24,66]). Unique peptide sequences or all MS/MS data can be used for homology search in publicly available and/or home-made databases (e.g. transcriptomes and EST libraries (see §3.1), virtually translated into the six reading frames). A common problem is that the obtained MS or MS/MS spectra are typically matched against publicly available protein

databases that in the case of most adhesive-producing organisms are nonexistent or derive from incomplete genome assemblies and annotations leading to deficient protein identification. For instance, protein isoforms arising from genetic polymorphisms, post-transcriptional events such as RNA-editing and PTMs will be largely missed (see [119] for review). This can be circumvented by coupling MS approaches with whole genome and total RNA sequencing, allowing the generation of near-complete databases of genetic variation and its transcripts for each adhesive organ (see §3.1). A dual proteomic and transcriptomic approach is therefore the best way leading to the identification of novel adhesive proteins and the retrieval of their complete sequences (e.g. the velvet worm protein Er\_P1–3 [10]; the sea star protein Sfp1 [8]).

When the sequence of the unknown protein cannot be obtained from databases, a PCR approach can be employed. This technique relies on the design of degenerate oligonucleotide primers on the basis of the peptide sequences (figure 3). These primers are used to screen cDNA libraries (e.g. the mussel protein Mefp-5 [29]) or on mRNA reverse-transcribed into cDNA (RT-PCR) (e.g. the barnacle protein cp52k [13]). In this strategy, primers designed on N-terminal peptides are more convenient as only sense degenerate primers are used in combination with antisense universal primers encoded in the vector used to build the cDNA library (e.g. Mefp-5 [29]) or encoded by a specific adapter incorporated during the RT (e.g. the tubeworm proteins Pc-1 and Pc-2 [34]; the slug proteins Sm40 and Sm85 [33]). Degenerate primers designed on internal peptides, however, have to be tested as sense or antisense. RT-PCR can also be performed using a sense degenerate primer designed on an N-terminal peptide and an antisense degenerate primer designed on an internal peptide (e.g. the zebra mussel protein Dpfp1 [23]), or using two primers designed on both extremity of a same peptide (e.g. the barnacle protein Mrcp-100k [14]). Discrete PCR products obtained following these strategies are subsequently cloned and sequenced. The generated partial sequences are usually completed by 3'- and 5'- rapid amplification of cDNA ends (RACE)-PCR using sequence specific primers (e.g. the barnacle protein Mrcp-19k [11]). The full-length sequence of the adhesive protein is then deduced from the open reading frame of the cDNA sequence which, in eukaryotes, starts with an ATG codon (coding for methionine) and ends with a stop codon.

### 3.3. Identification of post-translational modifications

The physico-chemical properties of many adhesive proteins are derived in part from their PTMs [87,120]. In the adhesive proteins described so far, these modifications are of three main types: glycosylation, hydroxylation and phosphorylation.

#### 3.3.1. Glycosylation

An impressive variety of carbohydrate–peptide linkages have been described from glycoproteins found in essentially all living organisms. These glycopeptide bonds can be classified in different groups, the most common being the *N*-glycosidic bonds in which the oligosaccharide is attached to the amide group of an asparagine residue within the consensus peptide sequence NXS/T, where X is any amino acid except proline; *O*-glycosidic bonds in which a mono- or oligosaccharide is attached to the hydroxyl group of a serine or threonine residue; and *C*-mannosyl bonds in which a mannose residue is attached

to a tryptophan residue through a C–C bond (see [121,122] for review). Glycoproteins have been implicated in attachment processes of fouling organisms such as barnacles and mussels. Examples include the freshwater mussel adhesive proteins Dpfp-1 and Dpfp-2 [65], which present extensive threonine or serine *O*-glycosylation. The marine mussel adhesive protein Pvfp-1 was shown to possess not only extensive threonine *O*-glycosylation but also tryptophan mannosylation [32,123]. The roles of the sugar residues in mussel proteins are still speculative, but have been proposed to increase conformational stability and enhancement of protein binding ability [122]. C-linked mannosylation would indeed render tryptophan more polar and solvent-accessible. It was also observed to block cleavage in the vicinity of modified residues by exo- and endoproteases, hence rendering proteins more resistant to degradation [32]. In barnacles, one cement protein, cp-52k, is known to be *N*-glycosylated [13].

The analysis of protein glycans is complicated by their vast variety and the large number of potential glycosylation combinations. Moreover, the same glycosylation site can be occupied by different glycans in different copies of a protein (see [122] for review). Despite the detection of glycoproteins in the adhesives of a wide range of organisms, little is still known about the composition of their carbohydrate fraction and whether these carbohydrates are covalently attached to the proteins or not. Below are some of the methods that have been applied to bioadhesives.

*Detection by staining.* A basic, simple method to determine whether a protein is glycosylated is to resolve it by one-dimensional or two-dimensional electrophoresis and to stain the gel for glycoproteins. Most gel-staining procedures are based on the periodic acid–Schiff reaction, in which periodic acid oxidizes vicinal diols in carbohydrate residues to form two aldehyde groups, which react with the Schiff reagent to give a magenta colour. This reaction was used by Ohkawa *et al.* [123] to demonstrate the presence of *O*-glycosylated threonines in Pvfp-1. Other reactions use periodate oxidation followed by biotinylation and binding of peroxidase-labelled avidin. This method was used by Kamino *et al.* [13] to highlight the *N*-glycosylation of cp-52k. Affinity-based staining can also be performed using lectins which are proteins that specifically bind mono- or oligosaccharides, allowing glycoprotein detection and characterization through lectin histochemistry and lectin blotting. Hennebert *et al.* [124] used a set of 16 lectins to label tube foot sections, footprints, and proteins extracted from these footprints in the sea star *A. rubens*. One *N*-linked and one *O*-linked glycoprotein were detected, their oligosaccharide chains enclosing galactose, *N*-acetylgalactosamine, fucose and sialic acid residues.

*Glycan structure analysis.* Once protein glycosylation has been confirmed, the glycan moiety structure can be further studied by liquid chromatography and MS. The glycan can be analysed either attached to the protein or following its release, the latter being more reliable. Depending on its nature, the glycan moiety can be removed from the protein by chemical or enzymatic means. For chemical removal, the two main methods for removing *O*- and *N*-linked oligosaccharides are  $\beta$ -elimination and hydrazinolysis, respectively. Amino acid analysis of Dpfp-1 and Dpfp-2 following removal of carbohydrate residues by reductive  $\beta$ -elimination allowed Rzepecki & Waite [65] to detect heavy loss of threonine and serine residues, suggesting that approximately 76% of the *O*-glycans in Dpfp-1, and some 90% in Dpfp-2, were attached

to threonine rather than serine. For enzymatic cleavage, most of the commercially available enzymes are specific for *N*-glycans, with fewer being available for *O*-glycans. In addition, given the diversity of *O*-glycans several enzymes might be required for the analysis of a single sample, and in most cases chemical removal is preferable. The most used *N*-glycan-cleaving enzymes are PNGase F and A, or endoglycosidases such as Endo-H. After its release, the glycan moiety can be analysed by chromatography using a fluorescent label to improve its detection. Glycan chromatography is often coupled with mass spectrometry for the elucidation of glycan sequence, branching and linkage [125]. MS can also be used to analyse the protein moiety from which the glycan was removed, in order to identify the amino acids that were formerly attached to the glycans. Tandem MS was used to demonstrate that, in mussel Pvfp-1, some tryptophan residues are C2-mannosylated [32].

### 3.3.2. Hydroxylation

During the chemical process of hydroxylation, an amino acid residue is modified by the attachment of at least one hydroxyl group (–OH). Hydroxylation of amino acid side chains in proteins is less common than other PTMs such as glycosylation or phosphorylation [126,127]. Protein hydroxylation is however an important process in marine bioadhesion. Up to now, four amino acids have been found to be hydroxylated in mussel and tubeworm adhesive proteins, giving rise to five modified residues. These residues seem to confer to marine proteins their exceptional adhesive properties. Indeed, they may contribute to make the protein more competitive with water by creating hydrogen bonds with surfaces [58,91,128]. DOPA is formed by the hydroxylation of tyrosine residues by a polyphenol oxidase (tyrosinase<sup>1</sup>). This modified residue was found in all the plaque proteins (fp-1 to fp-6) of mussels [58,63] and in the tubeworm proteins Pc-1 and Pc-2 [57]. 4-Hydroxyproline and 3,4-dihydroxyproline are the two hydroxylated derivatives of proline. They were detected in the decapeptide consensus repeat of the marine mussel protein Mefp-1 [3,129]. 4-Hydroxyarginine was found in the marine mussel protein Mefp-3 [91]. This protein has the ability to make conformational changes owing to its small size and the presence of glycine residues, a structural flexibility which enables the hydroxylated arginines to form hydrogen bonds with the substrate. 7-Hydroxytryptophan was found in the sequence of the protein Pvfp-1 from the mussel *P. viridis* [32]. In contrast to Mefp-1, only trace levels of DOPA were detected in Pvfp-1, what suggested the possibility that DOPA may be functionally replaced by hydroxytryptophan. Indeed, this modified amino acid resembles DOPA in having attributes that contribute to both cohesive and adsorptive interactions necessary for adhesion.

Below are some of the methods used to detect hydroxylation (mostly tyrosine hydroxylation) in bioadhesives.

*Amino acid analysis.* This method requires the hydrolysis of proteins, generally under acid conditions. Resulting amino acids are then separated by chromatography, detected and quantified by comparison with internal and external standards [130,131]. For instance, amino acid analysis allowed the highlighting of the presence of 3,4- and 4-hydroxyproline, and DOPA in mussel adhesive [132].

*Detection by staining.* Arnow staining specifically labels catechols, including DOPA [133]. The chemistry of the technique involves nitration of *o*-diphenols (catechols), resulting

in a relatively stable diphenolate derivative with strong red colour. The nitroblue tetrazolium method indicates the presence of redox active compounds. It is based on the ability of DOPA-containing proteins to reduce nitroblue tetrazolium in the presence of alkaline pH and an excess of glycine (used as reducing agent). The resulting product, formazan, is a deep blue coloured compound [134]. Both methods were used to visualize the distribution of DOPA-containing proteins in the cement cells of the tubeworms *P. californica* [34,56] and *S. alveolata* [36]. The same methods can also be used on gels to stain DOPA-containing proteins after they have been separated by electrophoresis (e.g. [57,111]).

**Mass spectrometry.** MS is generally used in combination with a wide variety of separation methods (gel electrophoresis and chromatography techniques) to detect hydroxylation (see §3.2 for full explanation about MS). For instance, this methodology has been widely used to detect hydroxylation of the arginine residues in the mussel protein Mefp-3 [91].

**Quantitative UV spectroscopy.** The proportion of tyrosine residues modified to DOPA can be quantitatively measured by UV spectroscopy. The absorbencies of a 1 mM DOPA standard and of the DOPA-containing protein can be measured at 250–350 nm using a UV–visible spectrophotometer. The spectra differences can be determined by subtracting the spectra of the 1 mM DOPA standard from that obtained from the modified sample. Using the  $\lambda_{\max}$  and  $\Delta\epsilon$  value of the 1 mM DOPA standard, the number of DOPA residues in the protein sample can be calculated according to the Lambert–Beer law [135].

### 3.3.3. Phosphorylation

Protein phosphorylation is an important regulator of both cellular and extracellular events. Recently, protein phosphorylation has also emerged as an important process in biological adhesives. Among the four types of phosphorylation currently described (*O*-phosphorylation, *N*-phosphorylation, *S*-phosphorylation and acylphosphorylation) [136,137], only *O*-phosphorylation of serine residues has been detected in some marine adhesive proteins [87,120]. So far phosphoproteins have been reported among mussel foot proteins (Mefp-5 [29], Mgfp-5 [138], Mcfp-5 and Mcfp-6 [26]), and tubeworm cement proteins (Pc-3A and Pc-B [34]). In marine adhesives, phosphorylation is thought to impart a potential for both cohesive (by  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  bridging) and adhesive (as an adaptation for adhesion to calcareous substrata) contributions to the adhesive [26,53,84,118,139]. Moreover, phosphoserine (pSer) residues may also be involved in protein–protein cross-linking as they are thought to condense with histidine residues to form histidinoalanine cross-links with the loss of phosphate [140].

Different analytical methods have been developed for the detection and quantification of *O*-phosphorylation [136,137]. Below are some examples.

**Amino acid analysis.** The same procedure described for the detection of hydroxylation, although with some adjustment, can be applied to detect phosphoserine residues (e.g. in Mefp-5 [29]).

**Detection by staining.** Currently, fluorescent dyes are becoming the method of choice to stain phosphorylated proteins directly in acrylamide gels. The sensitivity is high (ng scale) and it can be combined with a total protein stain, allowing protein phosphorylation levels and expression

levels to be monitored in the same gel. This approach was used by Santos *et al.* [109] to pinpoint putative adhesive proteins from the proteome of the tube feet in the sea urchin *P. lividus*. Within the total proteome, only 2% of the proteins were related with cell–cell or cell–substrate adhesion. Among those, one protein is homologous to nectin, a sea urchin secreted protein involved in embryonic cell–cell and cell–substrate adhesion. Two-dimensional gel staining with specific fluorescent stains revealed that in the tube foot nectin is present in eight isoforms of which five are simultaneously phosphorylated and glycosylated. Recent data confirmed that nectin is secreted in the sea urchin adhesive footprint (A. Toubarro *et al.* 2015, unpublished data). Although these fluorescent stains are optimized for staining sodium dodecylsulfate polyacrylamide gel electrophoresis gels, they can also be used in histochemistry on tissue sections, or on the secreted adhesive [141]. Nowadays, a large number of phospho-specific antibodies are commercially available, allowing pSer affinity-based staining. These antibodies can be used in western blot on extracted proteins and peptides, and also in immunohistochemistry on tissue sections. Antibodies can also be used to enrich and purify phosphorylated proteins and peptides. The use of an anti-pSer monoclonal antibody allowed the detection of phosphorylated proteins in the cement cells of the honeycomb worm *S. alveolata*, suggesting that, in this species, a polyphosphoprotein homologous to the protein Pc3 of *P. californica* would be also present [36,120].

**Mass spectrometry.** MS may be applied not only for detection of phosphorylation, but also for the identification of phosphorylation sites. Phosphorylation is detected by the analysis of mass spectra of trypsin-digested peptides that present a mass shift ( $m/z$  79.9 or neutral loss  $m/z$  80 or 98) in comparison with the theoretical peptide mass. This technique provides high speed and high sensitivity for detection of phosphorylation, but signals from phosphopeptides are generally weaker as compared with non-phosphorylated peptides and, therefore, it can be difficult to observe the signals from low-abundance phosphoproteins in the high background of abundant non-phosphorylated proteins. This can be overcome by enrichment of phosphoproteins or phosphopeptides prior to MS. Waite & Qin [29] used MS to analyse purified Mefp-5 and observed peak intervals of about 80 Da suggestive of the presence of phosphate or sulfate groups. This was further confirmed by subjecting Mefp-5 to an alkaline phosphatase treatment that produced MS spectra with progressive mass decreases in decrements of 80–90 Da, which is the mass change associated with dephosphorylation.

### 3.4. Validation of the adhesive function of the protein

For a single organism or adhesive organ, the use of transcriptomics and proteomics can generate long lists of putative adhesive proteins (e.g. [101,109]). Not all these proteins are actually involved in the adhesion process and, after selection of an interesting candidate, it is important, therefore, to validate its adhesive function. This is usually done by confirming the localization of the protein in the adhesive cells as well as in the secreted adhesive, by knocking down the expression of the protein and evaluating the resulting phenotype, and/or by measuring the physico-chemical properties of recombinantly expressed proteins.

### 3.4.1. Localization of the protein in adhesive cells and in the secreted adhesive

Genes and proteins identified by transcriptomics or proteomics need to be validated by corroborating their expression in the respective adhesive tissue or cell type.

Polyclonal antibodies represent a straightforward tool for detecting proteins in whole mount preparations or tissue sections [142,143]. They offer great sensitivity, are simple to handle, easy to store, and allow flexible applications (western blots, immunohistochemistry, immuno-TEM). Polyclonal antibodies are generated against one peptide or a complete protein. In general, the protein sequence (full-length or partial) is submitted to a commercial polyclonal antibody service provider. After bioinformatic analyses, a selection of 1–3 peptides is provided. These peptides will be chemically synthesized and used for immunization of two rabbits. Serum or affinity-purified antibodies will be delivered to the customer after about four months (but there are also rapid protocols). Disadvantages of polyclonal antibodies are the variability of different batches generated in different animals at different times, a higher potential for cross reactivity, and the limited amount of serum. Monoclonal antibodies, on the other hand, are quantitatively not restricted but require elaborate equipment and experience or are expensive when outsourced. Antibodies have been used in several model organisms to confirm the presence of putative adhesive proteins in the adhesive cells: e.g. the protein Sfp1 in sea stars [8], the polybasic proteins Pc-2, Pc-4 and Pc-5 in tubeworms [56], the protein Dpfp1 in zebra mussels [144], the proteins AgSF1 and AgSF2 in spiders [18] or the cement protein RIM36 in ticks [20]. In all these studies, the antibodies were also used to localize the proteins in the secreted adhesive, permitting one to investigate whether the protein is present at the interface (adhesive function) or in the bulk (cohesive function) of the adhesive layer.

If antibodies cannot be obtained (e.g. low immunogenicity of the protein like in the case of Pc-1 from tubeworms [56]), the localization of the corresponding mRNA can be visualized by *in situ* hybridization (ISH) [100,145]. By means of ISH a labelled RNA probe is produced and allowed to hybridize with the complementary mRNA. The labelled probe is then detected by antibody staining. In the flatworm *Macrostomum lignano*, ISH has been applied to study an adhesion-related intermediate-filament protein, *macif1* [101]. Expression of the adhesive proteins Pc-1 to Pc-5 of the tubeworm *P. californica* has been studied by ISH, showing that some of these proteins are localized in cement cells with homogeneous granules while the others are localized in cement cells with heterogeneous granules [56]. Foot-specific expression of adhesion-related genes has also been demonstrated in both marine [146,147] and freshwater mussels [148,149].

### 3.4.2. RNA interference

Understanding of the function of a gene or protein involved in bioadhesion is essential to elucidate the molecular foundation of the adhesion mechanism. RNA interference (RNAi) revolutionized the analysis of gene function by taking advantage of a cellular machinery that ultimately leads to the degradation of the mRNA of the gene of interest (e.g. [150,151]). Thereby, a loss-of-function phenotype is generated allowing conclusions to be drawn regarding the function of the target gene. The simplicity of application and the specificity of the knock-down by which the loss of gene function analysis can

be achieved have boosted the use of RNAi in the past decade for model and non-model organisms. The RNAi mechanism is complex but today the components of the RNAi pathway are well understood [150,151]. Briefly, *in vitro* generated dsRNA molecules from 150 to 800 bp corresponding to the gene of interest are provided externally by soaking or are applied by injection. The dsRNA molecules become degraded into RNA duplexes of 21–25 bp in length by the endonuclease Dicer. The resulting short double stranded RNA molecules are called ‘small interfering RNAs’ (siRNAs). Next, Argonaute proteins bind and subsequently unwind the siRNA duplex into single strands. The antisense strand (the guiding strand) of the duplex becomes then incorporated into the RNA-induced silencing complex (RISC) while the other strand (the passenger strand) gets degraded. In the next step, messenger RNAs (mRNAs) with the complementary sequence to the RISC-located guiding strand are loaded into the RISC. Upon association between the two RNA molecules the mRNA gets cleaved. Eventually, this post-transcriptional gene silencing process results in the depletion of target mRNA molecules. Ultimately, as no new protein can be made, a knock-down phenotype of the respective transcript will be obtained. The lack of the particular protein will provide insight into its function. With respect to bioadhesion, the involved protein components could be functionally analysed, and their contribution to the adhesion process could be studied in detail with this method. In *M. lignano*, RNAi knock-down of *macif1* resulted in a lack-of-adhesion phenotype [101]. To the best of our knowledge, this is however the only study so far in which RNAi has been used to validate the adhesive function of a protein in metazoans.

### 3.4.3. Recombinant proteins

To understand in detail the role of each protein component in a biological adhesive, it is necessary to obtain these components in large quantities. To bypass the problem of extracting adhesive proteins directly from animals, biomimetic molecules have been produced in the form of recombinant preparations of the adhesive proteins. This production has allowed the investigation of the essential functional characteristics of the proteins such as adsorption properties and adhesive properties on microscopic and macroscopic scales (e.g. [11,152,153]). Description of the methods allowing one to obtain these characteristics is out of the scope of this review but details can be found in [154,155].

To produce the heterologous adhesive proteins, different prokaryotic, eukaryotic and cell or tissue culture-based expression systems can be used. Common hosts include bacteria (*Escherichia coli*), yeasts (*Saccharomyces cerevisiae*), plants (tobacco) and mammals (goat, rabbit and mouse). Cell culture-based expression systems employ insect, plant and mammalian cells or tissue cultures instead of using the whole organism. All expression hosts have advantages and disadvantages that should be considered when making a choice. Some important variables for successful heterologous protein expression systems include: a codon usage compatible host; existence of promoter, transcriptional and translational regulators; optimal cultivation and protein purification methods; the ability to perform PTMs and to achieve correct folding of the recombinant protein; the possibility to provide high yields of the heterologous protein and easy scale-up process [27].

The recombinant DNA technology approach has been used mostly to produce large quantities of mussel adhesive proteins (fps) (see [42] for review). First attempts to obtain recombinant forms of the protein Mefp-1 started in the early 1990s [156,157]. However, these studies failed to obtain large amounts of recombinant proteins. In a later study, Lee *et al.* [158] expressed Mefp-1 fused to an *E. coli* signal peptide and thus optimized the expression of recombinant Mefp-1 in bacteria. Hwang *et al.* carried out two different studies with recombinant *M. galloprovincialis* foot protein 5 (rMgfp-5 [31]) and 3 (rMgfp-3 [152]). Both recombinant proteins were fused to a histidine tag and produced in the soluble fraction of *E. coli*. Even though recombinant mussel adhesive proteins were successfully produced and showed good performance in microscale adhesion tests, macroscale testing and large-scale applications were still not feasible. The main reasons included toxicity to the expression host, low expression levels and low solubility of the purified protein [27,159,160]. An additional problem related to the heterologous expression of adhesive proteins is that, generally, the recombinant proteins lack PTMs. These modifications include hydroxylation of proline, arginine and/or tyrosine residues. The production of functional recombinant mussel fps, therefore, requires an additional *in vitro* modification step: the enzyme-catalysed modification of tyrosine residues into DOPA. This is usually done using a commercially available mushroom tyrosinase [160].

In addition to mussel proteins, recombinant adhesive proteins were produced in *E. coli* from sequences obtained in barnacles, tubeworms, spiders and ticks. Two barnacle recombinant adhesive proteins, rMrsp-19k [11] and rMrsp-20k [15], have been successfully produced. The protein Sa-1 from the tubeworm *S. alveolata* was also successfully expressed and the recombinant protein, rSa-1, was mainly produced in the insoluble fraction of the bacteria, probably as inclusion bodies [59]. In spiders, fragments of the attachment disc glue silk fibroin 2 (PySp2) from *N. clavipes* [17] and the aggregate gland protein AgSF1 from *L. hesperus* [18] were produced and used for artificial spinning. Finally, in the tick *R. haemaphysaloides*, the cement protein RH50 was produced and used to immunize rabbits in order to challenge tick infestation [22].

## References

- Gorb S. 2008 Biological attachment devices: exploring nature's diversity for biomimetics. *Phil. Trans. R. Soc. A* **366**, 1557–1574. (doi:10.1098/rsta.2007.2172)
- Nachtigall W. 1974 *Biological mechanisms of attachment: the comparative morphology and bionengineering of organs for linkage, suction, and adhesion*. New York, NY: Springer.
- Waite JH. 1983 Adhesion in byssally attached bivalves. *Biol. Rev.* **58**, 209–231. (doi:10.1111/j.1469-185X.1983.tb00387.x)
- Smith AM, Callow JA. 2006 *Biological adhesives*. Berlin, Germany: Springer.
- von Byern J, Grunwald I. 2010 *Biological adhesive systems: from nature to technical and medical application*. Vienna, Austria: Springer.
- Santos R, Aldred N, Gorb S, Flammang P. 2013 *Biological and biomimetic adhesives: challenges and opportunities*. Cambridge, UK: RSC Publishing.
- Guerette PA *et al.* 2013 Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science. *Nat. Biotechnol.* **31**, 908–915. (doi:10.1038/nbt.2671)
- Hennebert E, Wattiez R, Demeuldre M, Ladurner P, Hwang DS, Waite JH, Flammang P. 2014 Sea star tenacity mediated by a protein that fragments, then aggregates. *Proc. Natl Acad. Sci. USA* **111**, 6317–6322. (doi:10.1073/pnas.1400089111)
- Jones I, Lindberg C, Jakobsson S, Hellqvist A, Hellman U, Borg B, Olsson PE. 2001 Molecular cloning and characterization of spiggin. An androgen-regulated extraorganismal adhesive with structural similarities to von Willebrand factor-related proteins. *J. Biol. Chem.* **276**, 17 857–17 863. (doi:10.1074/jbc.M101142200)
- Haritos VS, Niranjane A, Weisman S, Trueman HE, Srisantha A, Sutherland TD. 2010 Harnessing disorder: onychophorans use highly unstructured proteins, not silks, for prey capture. *Proc. R. Soc. B* **277**, 3255–3263. (doi:10.1098/rspb.2010.0604)
- Urushida Y, Nakano M, Matsuda S, Inoue N, Kanai S, Kitamura N, Nishino T, Kamino K. 2007 Identification and functional characterization of a novel barnacle cement protein. *FEBS J.* **274**, 4336–4346. (doi:10.1111/j.1742-4658.2007.05965.x)
- Kamino K. 2001 Novel barnacle underwater adhesive protein is a charged amino acid-rich protein constituted by a Cys-rich repetitive sequence. *Biochem. J.* **356**, 503–507. (doi:10.1042/0264-6021:3560503)
- Kamino K, Nakano M, Kanai S. 2012 Significance of the conformation of building blocks in curing of barnacle underwater adhesive. *FEBS J.* **279**, 1750–1760. (doi:10.1111/j.1742-4658.2012.08552.x)

## 4. Conclusion

Many metazoans rely on adhesive secretions to perform diverse functions. The diversity of biological adhesives is therefore huge and they can involve a large range of components with different functions and interactions. However, being mainly protein based, biological adhesives can be characterized by modern molecular tools, including the 'omics' approach. In recent years, the combined use of transcriptomics and proteomics has emerged as the best way leading to the identification of novel adhesive proteins and retrieval of their complete sequences. For a single organism, however, the use of these tools can generate long lists of putative adhesive proteins which therefore require additional experiments to validate their function. When the extraction of adhesive molecules is hindered by the amount of tissue available or by the insolubility of the material, alternative molecular tools exist to circumvent the problems at different levels from genes to secreted proteins. The wide range of experimental strategies compiled in this review are useful not only for natural adhesives but also for other type of biological materials.

**Funding statement.** This work was supported by COST Action TD0906 (<http://www.cost-bioadhesives.org/>). E.H. and B.M., and P.F. are, respectively, postdoctoral researchers and research director of the Fund for Scientific Research of Belgium (FNRS). This study is a contribution of the Centre Interuniversitaire de Biologie Marine (CIBIM). P.L. is supported by the Austrian Science Fund FWF P25404. R.S. was supported by Fundação para a Ciência e Tecnologia through a research contract of the Ciência 2008 programme and a project grant (PTDC/MAR/117360/2010).

## Endnote

<sup>1</sup>Polyphenol oxidases catalyse the *o*-hydroxylation of monophenols (e.g. tyrosine) to *o*-diphenols (e.g. DOPA) and the further oxidation of *o*-diphenols to *o*-quinones. Enzyme nomenclature differentiates between monophenol monooxygenases, traditionally called tyrosinases (EC 1.14.18.1), and catechol oxidases (EC 1.10.3.1). Tyrosinases can oxidize tyrosine to L-DOPA and L-DOPA to *o*-quinone. By contrast, catechol oxidases oxidize L-DOPA to *o*-quinone. The formation of *o*-quinone is responsible for the browning/tanning/curing processes that occur in plants, bacteria and animals [161].

14. Kamino K, Inoue K, Murayama T, Takamatsu N, Harayama S, Shizuri Y. 2000 Barnacle cement proteins. Importance of disulfide bonds in their insolubility. *J. Biol. Chem.* **275**, 27 360–27 365. (doi:10.1074/jbc.M910363199)
15. Mori Y, Urushida Y, Nakano M, Uchiyama S, Kamino K. 2007 Calcite-specific coupling protein in barnacle underwater cement. *FEBS J.* **274**, 6436–6446. (doi:10.1111/j.1742-4658.2007.06161.x)
16. Choreshe O, Bayarmagnai B, Lewis RV. 2009 Spider web glue. Two proteins expressed from opposite strands of the same DNA sequence. *Biomacromolecules* **10**, 2852–2856. (doi:10.1021/bm900681w)
17. Geurts P *et al.* 2010 Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. *Biomacromolecules* **11**, 3495–3503. (doi:10.1021/bm101002w)
18. Vasanthavada K *et al.* 2012 Spider glue proteins have distinct architectures compared with traditional spidroin family members. *J. Biol. Chem.* **287**, 35 986–35 999. (doi:10.1074/jbc.M112.399816)
19. Blasingame E *et al.* 2009 Pyriform spidroin 1, a novel member of the silk gene family that anchors dragline silk fibers in attachment discs of the black widow spider, *Latrodectus hesperus*. *J. Biol. Chem.* **284**, 29 097–29 108. (doi:10.1074/jbc.M109.021378)
20. Bishop R, Lambson B, Wells C, Pandit P, Osaso J, Nkonge C, Morzaria S, Musoke A, Nene V. 2002 A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory e-granules of the type III salivary gland acini, induces strong antibody responses in cattle. *Int. J. Parasitol.* **32**, 833–842. (doi:10.1016/S0020-7519(02)00027-9)
21. Trimmell AR, Hails RS, Nuttall PA. 2002 Dual action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens. *Vaccine* **20**, 3560–3568. (doi:10.1016/S0264-410X(02)00334-1)
22. Zhou J, Gong H, Zhou Y, Xuan X, Fujisaki K. 2006 Identification of a glycine-rich protein from the tick *Rhipicephalus haemaphysaloides* and evaluation of its vaccine potential against tick feeding. *Parasitol. Res.* **100**, 77–84. (doi:10.1007/s00436-006-0243-7)
23. Anderson K, Waite JH. 1998 A Major protein precursor of zebra mussel (*Dreissena polymorpha*) byssus: deduced sequence and significance. *Biol. Bull.* **194**, 150–160. (doi:10.2307/1543045)
24. Gantayet A, Ohana L, Sone ED. 2013 Byssal proteins of the freshwater zebra mussel, *Dreissena polymorpha*. *Biofouling* **29**, 77–85. (doi:10.1080/08927014.2012.746672)
25. Zhao H, Robertson NB, Jewhurst SA, Waite JH. 2006 Probing the adhesive footprints of *Mytilus californianus* byssus. *J. Biol. Chem.* **281**, 11 090–11 096. (doi:10.1074/jbc.M510792200)
26. Zhao H, Waite JH. 2006 Linking adhesive and structural proteins in the attachment plaque of *Mytilus californianus*. *J. Biol. Chem.* **281**, 26 150–26 158. (doi:10.1074/jbc.M604357200)
27. Silverman HG, Roberto FF. 2007 Understanding marine mussel adhesion. *Mar. Biotechnol.* **9**, 661–681. (doi:10.1007/s10126-007-9053-x)
28. Warner SC, Waite JH. 1999 Expression of multiple forms of an adhesive plaque protein in an individual mussel, *Mytilus edulis*. *Mar. Biol.* **134**, 729–734. (doi:10.1007/s002270050589)
29. Waite JH, Qin X. 2001 Polyphosphoprotein from the adhesive pads of *Mytilus edulis*. *Biochemistry* **40**, 2887–2893. (doi:10.1021/bi002718x)
30. Inoue K, Odo S. 1994 The adhesive protein cDNA of *Mytilus galloprovincialis* encodes decapeptide repeats but no hexapeptide motif. *Biol. Bull.* **186**, 349–355. (doi:10.2307/1542281)
31. Hwang DS, Yoo HJ, Jun JH, Moon WK, Cha HJ. 2004 Expression of functional recombinant mussel adhesive protein Mgfp-5 in *Escherichia coli*. *Appl. Environ. Microbiol.* **70**, 3352–3359. (doi:10.1128/AEM.70.6.3352-3359.2004)
32. Zhao H, Sagert J, Hwang DS, Waite JH. 2009 Glycosylated hydroxytryptophan in a mussel adhesive protein from *Perna viridis*. *J. Biol. Chem.* **284**, 23 344–23 352. (doi:10.1074/jbc.M109.022517)
33. Li D, Graham L. 2007 Epidermal secretions of terrestrial flatworms and slugs: *Lehmannia valentiana* mucus contains matrilin-like proteins. *Comp. Biochem. Phys. B* **148**, 231–244. (doi:10.1016/j.cbpb.2007.06.001)
34. Zhao H, Sun C, Stewart RJ, Waite JH. 2005 Cement proteins of the tube-building polychaete *Phragmatopoma californica*. *J. Biol. Chem.* **280**, 42 938–42 944. (doi:10.1074/jbc.M508457200)
35. Endrizzi BJ, Stewart RJ. 2009 Glueomics: an expression survey of the adhesive gland of the sandcastle worm. *J. Adhesion* **85**, 546–559. (doi:10.1080/00218460902996457)
36. Becker PT, Lambert A, Lejeune A, Lanterbecq D, Flammang P. 2012 Identification, characterization, and expression levels of putative adhesive proteins from the tube-dwelling polychaete *Sabellaria alveolata*. *Biol. Bull.* **223**, 217–225. (doi:10.2307/41759008)
37. Flammang P. 2006 Adhesive secretions in echinoderms: an overview. In *Biological adhesives* (eds AM Smith, JA Callow), pp. 183–206. Berlin, Germany: Springer.
38. Flammang P. 1996 Adhesion in echinoderms. In *Echinoderm studies* (eds M Jangoux, JM Lawrence), pp. 1–60. Rotterdam, The Netherlands: Balkema.
39. Flammang P, Michel A, Van Cauwenberge A, Alexandre H, Jangoux M. 1998 A study of the temporary adhesion of the podia in the sea star *Asterias rubens* (Echinodermata, Asteroidea) through their footprints. *J. Exp. Biol.* **201**, 2383–2395.
40. Hennebert E, Viville P, Lazzaroni R, Flammang P. 2008 Micro- and nanostructure of the adhesive material secreted by the tube feet of the sea star *Asterias rubens*. *J. Struct. Biol.* **164**, 108–118. (doi:10.1016/j.jsb.2008.06.007)
41. Santos R, da Costa G, Franco C, Gomes-Alves P, Flammang P, Coelho AV. 2009 First insights into the biochemistry of tube foot adhesive from the sea urchin *Paracentrotus lividus* (Echinoidea, Echinodermata). *Mar. Biotechnol.* **11**, 686–698. (doi:10.1007/s10126-009-9182-5)
42. Hennebert E, Wattiez R, Waite JH, Flammang P. 2012 Characterization of the protein fraction of the temporary adhesive secreted by the tube feet of the sea star *Asterias rubens*. *Biofouling*. **28**, 289–303. (doi:10.1080/08927014.2012.672645)
43. Walker G. 1981 The adhesion of barnacles. *J. Adhesion* **12**, 51–58. (doi:10.1080/00218468108071188)
44. Kamino K. 2008 Underwater adhesive of marine organisms as the vital link between biological science and material science. *Mar. Biotechnol.* **10**, 111–121. (doi:10.1007/s10126-007-9076-3)
45. Lacombe D. 1970 A comparative study of the cement glands in some balanid barnacles (Cirripedia, Balanidae). *Biol. Bull.* **139**, 164–179. (doi:10.2307/1540134)
46. Power AM, Klepal W, Zheden V, Jonker J, McEvilly P, von Byern J. 2010 Mechanisms of adhesion in adult barnacles. In *Biological adhesive systems: from nature to technical and medical application* (eds J von Byern, I Grunwald), pp. 153–168. Vienna, Austria: Springer.
47. Jonker JL, von Byern J, Flammang P, Klepal W, Power AM. 2012 Unusual adhesive production system in the barnacle *Lepas anatifera*: an ultrastructural and histochemical investigation. *J. Morphol.* **273**, 1377–1391. (doi:10.1002/jmor.20067)
48. Walker G. 1972 The biochemical composition of the cement of two barnacle species, *Balanus hameri* and *Balanus crenatus*. *J. Mar. Biol. Assoc. UK* **52**, 429–435. (doi:10.1017/S0025315400018786)
49. Kamino K. 2006 Barnacle underwater attachment. In *Biological adhesives* (eds AM Smith, JA Callow), pp. 145–166. Berlin, Germany: Springer.
50. Kamino K. 2010 Molecular design of barnacle cement in comparison with those of mussel and tubeworm. *J. Adhesion* **86**, 96–110. (doi:10.1080/00218460903418139)
51. Rouse G, Pleijel F. 2001 *Polychaetes*. Oxford, UK: Oxford University Press.
52. Vovelle J. 1965 The tube of *Sabellaria alveolata*. *Arch. Zool. Exp. Gen.* **106**, 1–187.
53. Stewart RJ, Weaver JC, Morse DE, Waite JH. 2004 The tube cement of *Phragmatopoma californica*: a solid foam. *J. Exp. Biol.* **207**, 4727–4734. (doi:10.1242/jeb.01330)
54. Stevens MJ, Steren RE, Vlamidir H, Stewart RJ. 2007 Multiscale structure of the underwater adhesive of *Phragmatopoma californica*: a nanostructured latex with a steep microporosity gradient. *Langmuir* **20**, 5045–5049. (doi:10.1021/la063765e)
55. Wang CS, Stewart RJ. 2013 Multipart copolyelectrolyte adhesive of the sandcastle worm, *Phragmatopoma californica* (Fewkes): catechol oxidase catalyzed curing through peptidyl-DOPA. *Biomacromolecules* **14**, 1607–1617. (doi:10.1021/bm400251k)
56. Wang CS, Stewart RJ. 2012 Localization of the bioadhesive precursors of the sandcastle worm, *Phragmatopoma*. *J. Exp. Biol.* **215**, 351–361. (doi:10.1242/jeb.065011)
57. Waite JH, Jensen RA, Morse DE. 1992 Cement precursor proteins of the reef-building polychaete

- Phragmatopoma californica* (Fewkes). *Biochemistry* **31**, 5733–5738. (doi:10.1021/bi00140a007)
58. Lee BP, Messersmith PB, Israelachvili JN, Waite JH. 2011 Mussel-inspired adhesives and coatings. *Ann. Rev. Mater. Res.* **41**, 99–132. (doi:10.1146/annurev-matsci-062910-100429)
59. Hennebert E, Maldonado B, Van de Weerd C, Demeuldre M, Richter K, Rischka K, Flammang P. 2015 From sand tube to test tube: the adhesive secretion of Sabellariid tubeworms. In *Bioadhesion and biomimetics: from nature to applications* (eds H Bianco-Peled, M Davidovich-Pinhas), pp. 105–123. Singapore: Pan Stanford Publishing.
60. Tamarin A, Lewis P, Askey J. 1976 The structure and formation of the byssus attachment plaque in *Mytilus*. *J. Morph.* **149**, 199–222. (doi:10.1002/jmor.1051490205)
61. Benedict CV, Waite JH. 1986 Composition and ultrastructure of the byssus of *Mytilus edulis*. *J. Morphol.* **189**, 261–270. (doi:10.1002/jmor.1051890305)
62. Waite JH. 2002 Adhesion à la moule. *Integr. Comp. Biol.* **42**, 1172–1180. (doi:10.1093/icb/42.6.1172)
63. Waite JH, Andersen NH, Jewhurst S, Sun C. 2005 Mussel adhesion: finding the tricks worth mimicking. *J. Adhesion* **81**, 297–317. (doi:10.1080/00218460590944602)
64. Farsad N, Sone ED. 2012 Zebra mussel adhesion: structure of the byssal adhesive apparatus in the freshwater mussel, *Dreissena polymorpha*. *J. Struct. Biol.* **177**, 613–620. (doi:10.1016/j.jsb.2012.01.011)
65. Rzepecki LM, Waite JH. 1993 The byssus of the zebra mussel, *Dreissena polymorpha*. II. Structure and polymorphism of byssal polyphenolic protein families. *Mol. Mar. Biol. Biotechnol.* **2**, 267–279.
66. Gantayet A, Ohana L, Sone ED. 2014 Novel proteins identified in the insoluble byssal matrix of the freshwater zebra mussel. *Mar. Biotechnol.* **16**, 144–155. (doi:10.1007/s10126-013-9537-9)
67. Barber I, Nettleship S. 2010 From 'trash fish' to supermodel: the rise and rise of the three-spined stickleback in evolution and ecology. *Biologist* **57**, 15–21.
68. Jakobsson S, Borg B, Haux C, Hyllner SJ. 1999 An 11-ketotestosterone induced kidney-secreted protein: the nest building glue from male three-spined stickleback, *Gasterosteus aculeatus*. *Fish Physiol. Biochem.* **20**, 79–85. (doi:10.1023/A:1007776016610)
69. Kawahara R, Nishida M. 2006 Multiple occurrences of spiggin genes in sticklebacks. *Gene* **373**, 58–66. (doi:10.1016/j.gene.2006.01.008)
70. Seear PJ, Head ML, Tilley CA, Rosato E, Barber I. 2014 Flow-mediated plasticity in the expression of stickleback nesting glue genes. *Ecol. Evol.* **4**, 1233–1242. (doi:10.1002/ece3.1016)
71. Walker G. 1987 Marine organisms and their adhesion. In *Synthetic adhesives and sealants* (ed. WC Wake), pp. 112–135. Chichester, UK: John Wiley & Sons.
72. Smith AM. 2006 The biochemistry and mechanics of gastropod adhesive gels. In *Biological adhesives* (eds AM Smith, JA Callow), pp. 167–182. Berlin, Germany: Springer.
73. Read VMJ, Hughes RN. 1987 Feeding behavior and prey choice in *Macropripatus torquatus* (Onychophora). *Proc. R. Soc. Lond. B* **230**, 483–506. (doi:10.1098/rspb.1987.0030)
74. Baer A, Mayer G. 2012 Comparative anatomy of slime glands in onychophora (velvet worms). *J. Morphol.* **273**, 1079–1088. (doi:10.1002/jmor.20044)
75. Benkendorf K, Beardmore K, Gooley AA, Packer NH, Tait NN. 1999 Characterisation of the slime gland secretion from the peripatus, *Euperipatoides kanangensis* (Onychophora: Peripatopsidae). *Comp. Biochem. Physiol. B* **124**, 457–465. (doi:10.1016/S0305-0491(99)00145-5)
76. Morera-Brenes B, Monge-Najera J. 2010 A new giant species of placentated worm and the mechanism by which onychophorans weave their nets (Onychophora: Peripatidae). *Rev. Biol. Trop.* **58**, 1127–1142. (doi:10.15517/rbt.v58i4.5398)
77. Perry DJ, Bittencourt D, Silberg-Liberles J, Rech EL, Lewis RV. 2010 Piriform spider silk sequences reveal unique repetitive elements. *Biomacromolecules* **11**, 3000–3006. (doi:10.1021/bm1007585)
78. Sahni V, Blackledge AT, Dhinojwala A. 2011 Changes in the adhesive properties of spider aggregate glue during the evolution of cobwebs. *Sci. Rep.* **1**, 41. (doi:10.1038/srep00041)
79. Sahni V, Miyoshi T, Chen K, Jain D, Blamires SJ, Blackledge TA, Dhinojwala A. 2014 Direct solvation of glycoproteins by salts in spider silk glues enhances adhesion and helps to explain the evolution of modern spider orb webs. *Biomacromolecules* **15**, 1225–1232. (doi:10.1021/bm401800y)
80. Maruyama SR *et al.* 2010 The expression of genes coding for distinct types of glycine-rich proteins varies according to the biology of three metastriate ticks, *Rhipicephalus* (*Boophilus*) *microplus*, *Rhipicephalus sanguineus* and *Amblyomma cajennense*. *BMC Genomics* **11**, 363. (doi:10.1186/1471-2164-11-363)
81. Harnnoi T, Watchabunsook S, Sakaguchi T, Xuan X, Fujisaki K. 2006 Characterization of *Haemaphysalis longicornis* recombinant cement-like antigens and preliminary study of their vaccination effects. *J. Vet. Med. Sci.* **68**, 1289–1295. (doi:10.1292/jvms.68.1289)
82. Anatriello I *et al.* 2010 An insight into the sialotranscriptome of the brown dog tick, *Rhipicephalus sanguineus*. *BMC Genomics* **11**, 450. (doi:10.1186/1471-2164-11-450)
83. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002 *Molecular biology of the cell*, 4th edn. New York, NY: Garland Science.
84. Stewart RJ, Wang CS, Shao H. 2011 Complex coacervates as a foundation for synthetic underwater adhesives. *Adv. Colloid Interface Sci.* **167**, 85–93. (doi:10.1016/j.cis.2010.10.009)
85. Wei JY, Broomell C, Israelachvili JN, Waite JH. 2013 Hydrophobic enhancement of Dopa-mediated adhesion in a mussel foot protein. *J. Am. Chem. Soc.* **135**, 377–383. (doi:10.1021/ja309590f)
86. Wei W, Tan Y, Martinez Rodriguez NR, Yu J, Israelachvili JN, Waite JH. 2014 A mussel-derived one component adhesive coacervate. *Acta Biomater.* **10**, 1663–1670. (doi:10.1016/j.actbio.2013.09.007)
87. Sagert J, Sun C, Waite H. 2006 Chemical subtleties of mussel and polychaete holdfasts. In *Biological adhesives* (eds AM Smith, JA Callow), pp. 125–143. Berlin, Germany: Springer.
88. Burzio LA, Waite JH. 2000 Cross-linking in adhesive quinoproteins: studies with model decapeptides. *Biochemistry* **39**, 11 147–11 153. (doi:10.1021/bi0002434)
89. Waite JH. 1985 Catechol oxidase in the byssus of the common mussel, *Mytilus edulis* L. *J. Mar. Biol. Assoc. UK* **65**, 359–371. (doi:10.1017/S0025315400050487)
90. Dickinson GH, Vega IE, Wahl KJ, Orihuela B, Beyley V, Rodriguez EN, Everett RK, Bonaventura J, Rittschof D. 2009 Barnacle cement: a polymerization model based on evolutionary concepts. *J. Exp. Biol.* **212**, 3499–3510. (doi:10.1242/jeb.029884)
91. Papov VV, Diamond TV, Biemann K, Waite JH. 1995 Hydroxyarginine containing polyphenolic proteins in the adhesive plaques of the marine mussel *Mytilus edulis*. *J. Biol. Chem.* **270**, 20 183–20 192. (doi:10.1074/jbc.270.34.20183)
92. Adams MD *et al.* 1991 Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651–1656. (doi:10.1126/science.2047873)
93. Strickler SR, Bombarely A, Mueller LA. 2012 Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *Am. J. Bot.* **99**, 257–266. (doi:10.3732/ajb.1200020)
94. Liu Y, Zhou J, White KP. 2013 RNA-seq differential expression studies: more sequence, or more replication? *Bioinformatics* **14**, 797. (doi:10.1093/bioinformatics/btt688)
95. Havlikova S, Roller L, Koci J, Trimmell AR, Kazimirova M, Klempa B, Nuttall PA. 2009 Functional role of 64P, the candidate transmission-blocking caccine antigen from the tick, *Rhipicephalus appendiculatus*. *Int. J. Parasitol.* **39**, 1485–1494. (doi:10.1016/j.ijpara.2009.05.005)
96. Oshlack A, Robinson MD, Young MD. 2010 From RNA-seq reads to differential expression results. *Genome Biol.* **11**, 220. (doi:10.1186/gb-2010-11-12-220)
97. Soneson C, Delorenzi M. 2013 A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinf.* **14**, 91. (doi:10.1186/1471-2105-14-91)
98. Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A. 2011 Differential expression in RNA-seq: a matter of depth. *Genome Res.* **21**, 2213–2223. (doi:10.1101/gr.124321.111)
99. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. 2013 Differential analysis of gene regulation at transcript resolution with

- RNA-seq. *Nat. Biotechnol.* **31**, 46–53. (doi:10.1038/nbt.2450)
100. Rodrigues ML, Lengerer B, Ostermann T, Ladurner P. 2014 Molecular biology approaches in bioadhesion research. *Beilstein J. Nanotechnol.* **5**, 983–993. (doi:10.3762/bjnano.5.112)
101. Lengerer B *et al.* 2014 Biological adhesion of the flatworm *Macrostomum lignano* relies on a duogland system and is mediated by a cell type-specific intermediate filament protein. *Front. Zool.* **11**, 12. (doi:10.1186/1742-9994-11-12)
102. Lin HC, Wong YH, Tsang LM, Chu KH, Qian PY, Chan BK. 2014 First study on gene expression of cement proteins and potential adhesion-related genes of a membranous-based barnacle as revealed from next-generation sequencing technology. *Biofouling* **30**, 169–181. (doi:10.1080/08927014.2013.853051)
103. Artimo P *et al.* 2012 ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**, W597–W603. (doi:10.1093/bioinformatics/bts133)
104. Nordahl Petersen T, Brunak S, von Heijne G, Nielsen H. 2011 SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786. (doi:10.1038/nmeth.1701)
105. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005 Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (ed. JM Walker), pp. 571–607. Totowa, NJ: Humana Press.
106. Heger A, Holm L. 2000 Rapid automatic detection and alignment of repeats in protein sequences. *Proteins* **41**, 224–237. (doi:10.1002/1097-0134(20001101)41:23.0.CO;2-Z)
107. Marchler-Bauer A *et al.* 2011 CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, D225–D229. (doi:10.1093/nar/gkq1189)
108. Waite JH, Hansen DC, Little KT. 1989 The glue protein of ribbed mussels (*Geukensia demissa*): a natural adhesive with some features of collagen. *J. Comp. Physiol. B* **159**, 517–525. (doi:10.1007/BF00694376)
109. Santos R, Barreto A, Franco C, Coelho AV. 2013 Mapping sea urchins tube feet proteome—a unique hydraulic mechano-sensory adhesive organ. *J. Proteomics* **79**, 100–113. (doi:10.1016/j.jprot.2012.12.004)
110. Zheden V, Klepal W, von Byern J, Bogner FR, Thiel K, Kowalik T, Grunwald I. 2014 Biochemical analyses of the cement float of the goose barnacle *Dosima fascicularis*—a preliminary study. *Biofouling* **30**, 949–963. (doi:10.1080/08927014.2014.954557)
111. Rzepecki LM, Hansen KM, Waite JH. 1992 Characterization of a cystine-rich polyphenolic protein family from the blue mussel *Mytilus edulis* L. *Biol. Bull.* **183**, 123–137. (doi:10.2307/1542413)
112. Anderson KE, Waite JH. 2002 Biochemical characterization of a byssal protein from *Dreissena bugensis* (Andrusov). *Biofouling* **18**, 37–45. (doi:10.1080/08927010290017716)
113. Steen H, Mann M. 2004 The abc's (and xyz's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711. (doi:10.1038/nrm1468)
114. Wehr T. 2006 Top-down versus bottom-up approaches in proteomics. *LC-GC North Am.* **24**, 1004.
115. Domon B, Aebersold R. 2006 Mass spectrometry and protein analysis. *Science* **312**, 212–217. (doi:10.1126/science.1124619)
116. Walter TC, Mann M. 2010 Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.* **190**, 491–500. (doi:10.1083/jcb.201004052)
117. Chevalier F. 2010 Highlights on the capacities of 'gel-based' proteomics. *Proteome Sci.* **8**, 23. (doi:10.1186/1477-5956-8-23)
118. NCBI Resource Coordinators. 2014 Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **42**, D7–D17. (doi:10.1093/nar/gkt1146)
119. Low TY *et al.* 2013 Quantitative and qualitative proteome characteristics extracted from in-depth integrated genomics and proteomics analysis. *Cell Rep.* **5**, 1469–1478. (doi:10.1016/j.celrep.2013.10.041)
120. Flammang P, Lambert A, Bailly P, Hennebert E. 2009 Polyphosphoprotein-containing marine adhesives. *J. Adhesion* **85**, 447–464. (doi:10.1080/00218460902996358)
121. Spiro RG. 2002 Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R–56R. (doi:10.1093/glycob/12.4.43R)
122. Roth Z, Yehezkel G, Khalaila I. 2012 Identification and quantification of protein glycosylation. *Int. J. Carbohydr. Chem.* **2012**, 640923. (doi:10.1155/2012/640923)
123. Ohkawa K, Nishida A, Yamamoto H, Waite JH. 2004 A glycosylated byssal precursor protein from the green mussel *Perna viridis* with modified dopa side-chains. *Biofouling* **20**, 101–115. (doi:10.1080/08927010410001681246)
124. Hennebert E, Wattiez R, Flammang P. 2011 Characterisation of the carbohydrate fraction of the temporary adhesive secreted by the tube feet of the sea star *Asterias rubens*. *Mar. Biotechnol.* **13**, 484–495. (doi:10.1007/s10126-010-9319-6)
125. Pagett HE *et al.* 2012 Structural characterisation of the N-glycan moiety of the barnacle settlement-inducing protein complex (SIPC). *J. Exp. Biol.* **215**, 1192–1198. (doi:10.1242/jeb.063503)
126. Kaelin WG. 2005 Proline hydroxylation and gene expression. *Annu. Rev. Biochem.* **74**, 115–128. (doi:10.1146/annurev.biochem.74.082803.133142)
127. Chopra RK, Ananthanarayanan VS. 1982 Conformational implications of enzymatic proline hydroxylation in collagen. *Proc. Natl Acad. Sci. USA* **79**, 7180–7184. (doi:10.1073/pnas.79.23.7180)
128. Burzio LO, Burzio VA, Silva T, Burzio LA, Pardo J. 1997 Environmental bioadhesion: themes and applications. *Curr. Opin. Biotechnol.* **8**, 309–312. (doi:10.1016/S0958-1669(97)80008-0)
129. Taylor SW, Waite JH. 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**, 10 803–10 804. (doi:10.1021/ja00102a063)
130. Davidson I. 2003 Hydrolysis of samples for amino acid analysis. In *Methods in molecular biology*, vol. 211: *protein sequencing protocols* (ed. BJ Smith), pp. 111–122. Totwa, NJ: Humana Press Inc.
131. Macchi FD, Shen FJ, Keck RG, Harris RJ. 2000 Amino acid analysis, using postcolumn ninhydrin detection, in a biotechnology laboratory. In *Methods in molecular biology*, vol. 159: *amino acid analysis protocols* (eds C Cooper, N Packer, K Williams), pp. 9–30. Totwa, NJ: Humana Press Inc.
132. Waite JH, Tanzer M. 1981 Polyphenolic substance of *Mytilus edulis*: novel adhesive containing L-dopa and hydroxyproline. *Science* **212**, 1038–1040. (doi:10.1126/science.212.4498.1038)
133. Arnow LE. 1937 Colorimetric determination of components of 3,4-dihydroxyphenyl-L-alanine/tyrosine mixtures. *J. Biol. Chem.* **118**, 531–537.
134. Paz MA, Flückiger R, Boak A, Kagan HM, Gallop PM. 1991 Specific detection of quinoproteins by redox-cycling staining. *J. Biol. Chem.* **266**, 689–692.
135. Waite JH. 1984 Determination of (catecholato)-borate complex using difference spectrometry. *Anal. Chem.* **56**, 1935–1939. (doi:10.1021/ac00275a040)
136. Reinders J, Sickmann A. 2005 State-of-the-art in phosphoproteomics. *Proteomics* **5**, 4052–4061. (doi:10.1002/pmic.200401289)
137. Jia M, Lin KW, Souchelnytskyi S. 2012 Phosphoproteomics: detection, identification and importance of protein phosphorylation. In *Integrative proteomics* (eds HCE Leung, TK Man, RJ Flores). Rijeka, Croatia: In Tech. (doi:10.5772/31305)
138. Wang YJ, Zheng X, Zhang LH, Ohta Y. 2004 Cloning and sequencing of the gene encoding mussel adhesive protein from *Mytilus* sp JHX-2002. *Process Biochem.* **39**, 659–664. (doi:10.1016/S0032-9592(03)00165-1)
139. Sun C, Fantner GE, Adams J, Hansma PK, Waite JH. 2007 The role of calcium and magnesium in the concrete tubes of the sandcastle worm. *J. Exp. Biol.* **210**, 1481–1488. (doi:10.1242/jeb.02759)
140. Taylor CM, Wang W. 2007 Histidinoalanine: a crosslinking amino acid. *Tetrahedron* **63**, 9033–9047. (doi:10.1016/j.tet.2007.05.114)
141. Gohad NV *et al.* 2014 Synergistic roles for lipids and proteins in the permanent adhesive of barnacle larvae. *Nat. Commun.* **5**, 4414. (doi:10.1038/ncomms5414)
142. Hendriksen CF. 2005 Introduction: laboratory animals and immunization procedures: challenges and opportunities. *ILAR J. NLMID* **46**, 227–229. (doi:10.1093/ilar.46.3.227)
143. Leenaars M, Hendriksen CF. 2005 Critical steps in the production of polyclonal and monoclonal antibodies: evaluation and recommendations. *ILAR J. NLMID* **46**, 269–279. (doi:10.1093/ilar.46.3.269)
144. Anderson KE, Waite JH. 2000 Immunolocalization of Dpfp1, a byssal protein of the zebra mussel *Dreissena polymorpha*. *J. Exp. Biol.* **203**, 3065–3076.

145. Wilcox JN. 1993 Fundamental principles of in situ hybridization. *J. Histochem. Cytochem.* **41**, 1725–1733. (doi:10.1177/41.12)
146. Inoue K, Takeuchi Y, Miki D, Odo S, Harayama S, Waite JH. 1996 Cloning, sequencing and sites of expression of genes for the hydroxyarginine-containing adhesive-plaque protein of the mussel *Mytilus galloprovincialis*. *Eur. J. Biochem.* **239**, 172–176. (doi:10.1111/j.1432-1033.1996.0172u.x)
147. Miki D, Takeuchi Y, Inoue K, Odo S. 1996 Expression sites of two byssal protein genes of *Mytilus galloprovincialis*. *Biol. Bull.* **190**, 213–217. (doi:10.2307/1542541)
148. Xu W, Faisal M. 2010 Factorial microarray analysis of zebra mussel (*Dreissena polymorpha*: Dreissenidae, Bivalvia) adhesion. *BMC genomics* **11**, 341. (doi:10.1186/1471-2164-11-341)
149. Xu W, Faisal M. 2010 Gene expression profiling during the byssogenesis of zebra mussel (*Dreissena polymorpha*). *Mol. Genet. Genomics* **283**, 327–339. (doi:10.1007/s00438-010-0517-8)
150. Agrawal N, Dasaradhi PV, Mohammed A, Malhotra P, Bhatnagar RK, Mukherjee SK. 2003 RNA interference: biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* **67**, 657–685. (doi:10.1128/MMBR.67.4.657-685.2003)
151. Wilson RC, Doudna JA. 2013 Molecular mechanisms of RNA interference. *Ann. Rev. Biophys.* **42**, 217–239. (10.1146/annurev-biophys-083012-130404)
152. Hwang DS, Gim Y, Cha HJ. 2005 Expression of functional recombinant mussel adhesive protein type 3A in *Escherichia coli*. *Biotechnol. Prog.* **21**, 965–970. (doi:10.1021/bp050014e)
153. Cha HJ, Hwang DS, Lim S, White JD, Matos-Perez CR, Wilker JJ. 2009 Bulk adhesive strength of recombinant hybrid mussel adhesive protein. *Biofouling* **25**, 99–107. (doi:10.1080/08927010802563108)
154. Berglin M, Hedlund J, Fant C, Elwing H. 2005 Use of surface-sensitive methods for the study of adsorption and cross-linking of marine bioadhesives. *J. Adhesion* **81**, 805–822. (doi:10.1080/00218460500189059)
155. Petrone L. 2013 Molecular surface chemistry in marine bioadhesion. *Adv. Colloid Interface Sci.* **195–196**, 1–18. (doi:10.1016/j.cis.2013.03.006)
156. Filpula DR, Lee SM, Link RP, Strausberg SL, Strausberg RL. 1990 Structural and functional repetition in a marine mussel adhesive protein. *Biotechnol. Progr.* **6**, 171–177. (doi:10.1021/bp00003a001)
157. Salerno AJ, Goldberg I. 1993 Cloning, expression, and characterization of a synthetic analog to the bioadhesive precursor protein of the sea mussel *Mytilus edulis*. *Appl. Microbiol. Biotechnol.* **39**, 221–226. (doi:10.1007/BF00228610)
158. Lee SJ, Han YH, Nam BH, Kim YO, Reeves P. 2008 A novel expression system for recombinant marine mussel adhesive protein Mefp1 using a truncated OmpA signal peptide. *Mol. Cells* **26**, 34–40.
159. Cha HJ, Hwang DS, Lim S. 2008 Development of bioadhesives from marine mussels. *Biotechnol. J.* **3**, 631–638. (doi:10.1002/biot.200700258)
160. Stewart RJ. 2011 Protein-based underwater adhesives and the prospects for their biotechnological production. *Appl. Microbiol. Biotechnol.* **89**, 27–33. (doi:10.1007/s00253-010-2913-8)
161. Waite JH. 1990 The phylogeny and chemical diversity of quinone-tanned glues and varnishes. *Comp. Physiol. Biochem. B* **97**, 19–29. (doi:10.1016/0305-0491(90)90172-P)