CHAPTER 4

BIOCHEMICAL AND SPECTRAL ANALYSES OF DIFFERENT RECOMBINANT GFP-LIKE PROTEINS FROM THE ACROPORA PRESETTLEMENT cDNA LIBRARY

4.1 Introduction

GFP-like proteins represent a diverse group of molecules found mainly in Hydrozoa and Anthozoa cnidarians and a few marine bilaterians (Shagin et al. 2004, Masuda et al. 2006, Mocz 2007). This diversity can be broadly categorized into different fluorescent (green, yellow, and red) and non-fluorescent (purple-blue) colour phenotypes (Labas et al. 2002). The first GFP-like protein was isolated and cloned from the bioluminescent hydrozoan jellyfish *Aquorea victoria* (Prasher et al. 1992, Prasher 1995, Niwa et al. 1996, Ormo et al. 1996). In the jellyfish, the GFP (avGFP) is coupled with a photoprotein called Aequorin and is bioluminescent. The photoprotein binds a luciferin-like molecule called coelenterezine that, upon oxidation, releases a photon. The oxidation of coelenterezine is mediated by conformational changes of Aequorin that are induced by the binding of Ca$^{2+}$ ions released by the cell. The light produced from this reaction is blue (490 nm); however, when photons are absorbed by the avGFP chromophore (p-hydroxybenzyl dineimidazolidinone), the phenoxide group acts as a donor to the imidazolidinone acceptor, and, by a charge-transfer mechanism, light is reemitted as fluorescence in the green spectral region (Morin 1974).

The GFP-like protein is composed of 11 β-sheets that fold to form a cylinder 30 Å in circumference and 40 Å in height (Tsien 1998, Kummer 2000). An α-helix bearing the tripeptide “XYG” motif (positions 63–65) runs through the centre of the cylinder and is responsible for the fluorescent properties of the protein. The tertiary arrangement of the protein is responsible for its green fluorescence. A series of torsional adjustments that occur during protein folding relocate the carboxyl carbon of one residue close to the amino group of the other; then, by oxygen-dependent cyclization and dehydrogenation reactions, an imidazolin-5-one heterocyclic ring systems is formed.
(Cubitt et al. 1995, Barondeau et al. 2003). For the red fluorescence to occur (e.g., in corallimorpharian *Discosoma DsRed*), an additional oxygen-dependent dehydrogenation that extends the “green-like” structure by two electron-withdrawing double bonds is required (Baird et al. 2000, Fradkov et al. 2000, Gross et al. 2000). In certain scleractinean GFP-like proteins, the red fluorescence is acquired by photon-mediated fragmentation of the protein backbone, resulting in the acquisition of the extended red-like chromophore (Pakhomov et al. 2004, Oswald et al. 2007). The behaviour of these chromophores allows for the tracking of desired objects in the cell. For example, cellular organelles labelled with GFP-like proteins that are exposed to ultraviolet light emit in the red region of the spectrum (Lukyanov et al. 2005). The yellow fluorescent protein (FP), cloned from a sea anemone (*Zoanthus* sp.), has a chromophoric arrangement in which a transiently appearing DsRed-like acylamine reacts with the terminal group of L66 to form a new six-membered ring that accounts for the yellow shift (Remington et al. 2005). Reactions similar to those that occur with the red chromophore also take place in non-fluorescent GFP-like chromoproteins (CPs); however, the non-coplanar conformation of the light-absorbing molecule results in an inability of the protein to emit light (Fradkov et al. 2000, Dove et al. 2001, Bulina et al. 2002, Verkhusha et al. 2004, Chan et al. 2006). Nevertheless, intense green light can still promote red fluorescence from supposed non-fluorescent CPs. CPs with the capability to fluoresce upon radiation are called kindling FPs (Chudakov et al. 2003).

The protein barrel is often a monomer in hydrozoan GFP-like molecules; however, most non-bioluminescent anthozoan GFPs are tetrameric. Their aggregation behaviour is mediated by the presence of two (hydrophobic and hydrophilic) interfaces formed by the interaction of the external side chains of different monomers (Dove et al. 2001, Campbell et al. 2002, Martynov et al. 2003, Gurskaya et al. 2006). In the biological context, aggregation has been related to light-trapping mechanisms, such as in reef organisms exposed to intense sunlight (Salih et al. 2004, Andresen et al. 2005). In the laboratory, however, this natural tendency for oligomerization represents a problem when the proteins are to be used as a tag (Campbell et al. 2002, Gurskaya et al. 2006).

Although much is known about GFP-like proteins in adult coral, there are no data concerning the actions and roles of these proteins in the early developmental stages. This is important because age-dependent changes in gene expression and GFP-like
protein composition could relate to the acquisition of symbionts in the free-living stage and in metamorphosis.

This chapter discusses the cloning and characterization of some GFP-like proteins from Acropora in the pre- and postsettlement stages. The proteins were characterized by their spectral properties (i.e., absorbance, excitation, and emission), molecular masses, and responses to changes in pH. Data acquired from these studies were compared to information found in the literature for all GFP-like molecules, and their similarities and differences are discussed.

4.1.1 Statement of Goals
By characterising several GFP-like variants it is possible to explain at the molecular level how certain mutations could be reflected into some measurable phenotypic characteristic such as differences in optical or solubility properties. With such information it is possible in principle not only to estimate the potential colour variation in the species, but also to explore the link between (non-spectroscopic) properties of the proteins and their possible roles in the biology of the coral. With respect to colour variation, the present study focussed mainly on larval material. Despite the fact that some of the proteins expressed in adult corals are spectroscopically similar to those expressed in larvae, the amino acid sequences are quite different but the functional significance of this is unclear. In addition, some of the mutant recombinant proteins produced had altered solubility relative to the parental protein. Some of the natural variants detected apparently differed with respect to oligomerization pattern. Other mutations not clearly implicated in spectroscopic or aggregative properties were analysed in a search for protein-protein binding patches potentially involved in heterologous protein-protein interactions.

Pure protein preparations were subjected to pH gradients in order to investigate the chemical characteristics of the chromophore. This information is critical for understanding the nature of the chromophore in the *A.millepora* proteins, and enables comparison with published information. This study also sheds some light on biophysical aspects of the GFP-like proteins and may have biotechnological applications.
4.2 Results

Results pertaining to the diversity of the expressed sequence tag (EST) clones present in the larval cDNA library and the phenotypic variations associated with the fluorescence or colour of each protein are found in this section. Three-dimensional (3D) models of some of the proteins were used to understand how “hot spots” in protein structure confer spectral characteristics of GFP-like proteins. In addition, characteristics such as molecular mass and response to pH are described for each cloned GFP-like protein.

4.2.1 GFP-like protein diversity observed in larvae ESTs

Information on the abundance/diversity of the GFP-like EST sequences was derived from an extensive array of up to 9000 sequences from three developmentally important stages: the prawn chip, planula or presettlement stage, and postsettlement stages that occur just after metamorphosis. Most of the sequences were derived from the planula stage library, although others came from the postsettlement library. No complete transcripts were detected at the prawn chip stage. All detected GFP-like sequences were grouped into two main contigs according to their nucleotide sequence composition. No other contigs containing a GFP-like sequence were observed in the EST collection. After visual comparison, each of the two contigs was revealed to contain fluorescent-like and non-fluorescent, CP-like genes.

The primary structures of the CPs recorded in the EST collections are compared in Figure 4.1. The non-fluorescent CP contig record showed 17 transcripts that coded for proteins 221 amino acids in length. Amino acid alignments of the different CPs were performed relative to the most abundant clone. On top of each line, the specific mutation is denoted with the relative proportion in parentheses. The alignment showed an overall variation of 5% in this molecule, with only 2–12 variable substitutions.

Figure 4.2 shows the primary structures of FPs. The FP contig contained a total of 31 clones, each with different frequencies. The length of the transcripts in all cases was 231 amino acids. The analysis was performed relative to the most abundant clone so that the substitutions were easy to detect. On the top of each line, the specific amino acid substitution is presented, and in parentheses is the proportion of that substitution. The amino acids responsible for fluorophore formation or those
involved in final autocatalytic fluorescence/colour development are highlighted. The alignment showed a 6% variation in the molecule, with only 4 of the 13 variable positions important for FP proportion.
Figure 4.1 Amino acid alignments of all complete clones that belong to the Chromo-like contig.
The names on the left represent the assigned identification for the EST clone and also include two bacterially expressed CPs that have distinctive absorbance maxima. The arrows point to the substitutions. Information concerning the proportion of each amino acid is shown in parentheses, and the number in the polypeptide chain is shown in bold type. The alignment was made relative to the clone amilGFP597, which was most abundant in the contig.
Figure 4.2 Amino acid alignments of all complete clones that belong to the Fluorescent-like contig.
The names on the left represent the assigned identification for the EST clone. Also included are two bacterially expressed CPs with distinctive absorbance maxima. The arrows point to the substitutions. Information concerning the proportion of each amino acid is in parentheses, and the number in the polypeptide chain is in bold type. The alignment was performed relative to the clone amilGFP597, which was the most abundant in the contig.
4.2.2 3D representation of the observed diversity in each contig

Several crystals of FPs have been produced, allowing for the construction of internet sites such as http://www.cbs.dtu.dk/services/CPHmodels/ that map atomic 3D arrangements of FPs from submitted amino acid sequences. For GFP-like proteins, all crystallized proteins have been shown thus far to possess identical β-can structures. Hence, it is reasonable to assume that the predictions made by the program are correct. The program, however, is incapable of reproducing the actual shape of the chromophore inside the β-barrel; the linear sequence of the three peptides that constitute the fluorophore are modified post-translationally, which regulates the tertiary structure of the protein to form the planar chromophore structure. Nevertheless, the models are important because they allow for comparisons of protein shapes and the distribution of specific residues of GFP-like proteins. In this study, the online engine was used to retrieve the protein model in .pdb format using Swiss-Pdb Viewer software (GlaxoSmithKline), and the appropriate modifications to the model were made.

Figure 4.3 shows the sequences of two phenotypically different clones, AmilCP597 and AmilCP601. Also shown is the total variation inside the CP contig from other non-expressed GFP-like ESTs. Both proteins display differences in the consecutive dipeptide N142H and T143S. The simulation shows that these substitutions are located at similar locations with respect to the chromophore; however, only the side chain that faces the internal β-can structure is in position N142H. Another substitution, V26G, is visible at the opposite side of the cylinder. The side chain of this amino acid faces the exterior. In the basal portion of the α-helix, a S53A substitution is observed in the two proteins. A substitution not recorded in these phenotypes but that was part of the contig variation is F10Y, which is a moderately abundant substitution. V37E is located close to the N-C termini, near the loop between two β-sheets and the side chain that faces outside the protein core.

Three variable sites (V92I, R101C, and S104T) were observed in the EST clones opposite to this substitution and just below the chromophore. Only R101C faces inward; the other two substitutions expose the side chain to the exterior. In this latter protein region, but closer to the N-C termini, the R151G and L154I substitutions expose their side chains to the exterior. Finally, the substitution L206S, located close to the base of the last β-sheet, faces outside the core of the protein.
Figure 4.4 includes the sequences of three FP clones categorized into two main phenotypes: the AmiGFP515a, b and the AmiRFP610. The figure also shows the diversity observed in the entire FP contig for the RFP model. The first position (close to the N-terminus) with variation was R10K; the side chain faces outside the Beta-can and is distributed between green and red variants. A13T, located in proximity to the last substitution, is observed in one of the green phenotypes (Figure 4.4B). The side chain faces the exterior of the protein. A49V was also a substitution observed in the green phenotype. The substitution faces outside the protein core; it has no apparent influence in the fluorescence phenotype because V49 is distributed in both the red and green variants.

Perhaps one of the most important substitutions for chromophore formation or variation is D66T, located just in the centre of the molecule as a part of the tripeptide motif. T66 was observed in both green FPs, and D66 was found only in the red FP. A166P was observed only in the RFP (Figure 4.4C); both GFPs have P166, although the side chain faces the exterior of the protein. R180Q and S183T are two substitutions close to the 166 position. R180 was observed only in Ami515b; the other green phenotype and the red phenotype carry Q180. The side chain faces outside the protein core.

Although S183T was present in both green phenotypes, the red phenotypes showed T183 at this position. The side chain faces the interior of the protein. Judging by its relative closeness to the chromophore, it is possible that these positions have some effect on the fluorescence phenotype. The rest of the substitutions were minor (in abundance) and were not expressed in the bacteria-based system. In the figure, however, the position on the RFP model represents the overall variation in the molecule.

The N11D substitution occurred in only one clone in the proximity of the first major variable site, R10K, which was between both coloured phenotypes. The side chain faces the exterior of the protein. E101D was also represented by one clone; it was located at the base of the cylinder, opposite to the N-C termini, with the side chain facing the interior of the protein.

Q127H was observed in two clones, presumably belonging to the red-like FPs. The substitution G134D, observed in one clone, was located in the middle of one of the
lids of the β-sheet barrel opposite the N-C termini. The amino acid side chain faces the exterior of the protein cylinder.

Another isolated substitution is S220A, which was located near the C-terminus. It is part of the last β-sheet and faces outside of the protein. The last amino acid of the protein also showed variability; P/V231A substitutions occurred, with P231 representing 16% and V231 representing 3% of the population. These amino acids do not appear to form part of the cylinder body.
Figure 4.3 3D model of the CP recombinant proteins and the substitutions observed in the whole contig. A–D and E–H represent amilCP601 and amilCP597, respectively, viewed laterally and from the top. The amino acids indicated by arrows are the substitutions that were compared with the other clone. Other amino acids that are not indicated by arrows represent the overall variation observed in the contig.
Figure 4.4 3D model representation of fluorescent recombinant proteins, including the observed diversity of amino acid substitutions inside the contig.

A-A', B-B', and C-C' represent the side and top views of the protein from the AmiGFP517a (clone 52c2), AmiGFP517b (clone c012c9), and AmiRFP602 (clone c018g5) clones, respectively. The amino acids indicated by the arrows represent the substitutions that were compared with the other clones. Other amino acids that are not indicated by arrows represent the variation observed in the whole contig.
4.2.3 Molecular masses of the proteins

To confirm the theoretical molecular mass of the proteins predicted by the amino acids sequence, identical amounts of the recombinant proteins were separated by 13% SDS-PAGE, and the proteins were visualized after Coomassie Blue staining. For AmilGFP517b (EST clone c012-c9), a major band at 29 kDa was flanked by two minor bands at 26 and 35 kDa (Figure 4.5A, lane 1). The AmilGFP517a (clone 114c8) banding signature (Figure 4.5A, lane 3) was almost identical to its sibling green fluorescent clone; however, the 26 kDa band was not observed for this recombinant protein. The AmilRFP602 protein migrated to the 29 kDa position (Figure 4.5A, lane 2), and a faint band was observed at ~10–12 kDa. The AmilCP597 and AmilCP601 CPs (Figure 4.5A, lanes 4 and 5, respectively) migrated primarily to 27 kDa; three faint bands at 47, 90, and 120 kDa were also observed.
**Figure 4.5 Phenotypic comparison of the expressed recombinant GFP-like proteins.**

A) Coomassie blue–stained GFP-like recombinant proteins after separation on 13% SDS polyacrylamide gels: 1) Amil517b (EST clone c012c9); 2) AmilRFP602 (EST clone c018-g5); 3) AmilGFP517a (clone 114c8); 4) AmilCP597 (clone 2;2); 5) AmilCP601 (EST Clone c015-a8). The relative molecular mass is displayed at the left of each figure in daltons. B) Visual comparison under normal light conditions of the eluted pure proteins in solution. Also shown is the bacterial culture expressing each protein excited with UV light: 1 and 5) AmilGFP517 under normal and UV light; 2 and 6) AmilRFP602 under normal and UV light; 3 and 4) AmilCP597 and AmilCP601 observed with normal light; C) *E. coli* expressing the recombinant proteins. 1–3) AmilGFP, RFP, and CP, respectively; D) Fluorescence of the CP after being blotted on 3M paper. 1–3) CP observed under normal light; the GREEN filter (designed to observe red fluorescence) and GFP2 filter (designed to observe the green fluorescence) were used.
4.2.4 Spectral characterization of GFP-like recombinant proteins from *Acropora millepora*

This section deals with the spectral signature of the recombinant proteins in solution. Scans were taken across the visible region to obtain absorbance, excitation, and emission signals. In all graphic representations, the X axis represents the wavelength in nanometres, and the principal Y axis represents the value in absorbance and intensity (excitation and emission) relative to the main peak. The secondary Y axis is the fourth derivative analyses of the data (absorbance, excitation, and emission in arbitrary units) with respect to the wavelength. This line overlaps the spectral data and can be used to identify the internal peaks under the curve.

4.2.4.1 Spectral signature of *A. millepora* green FP (AmilGFP517a, b)

Two different proteins that had green fluorescence were cloned. Despite the amino acid differences between them, they had identical emission characteristics. However, one of the clones (114c8, amilGFP517a) showed distinct hydrophobic characteristics during purification. After clarification of the cell homogenate, it was difficult to suspend AmilGFP517a in the aqueous phase for separation by affinity chromatography.

The green FPs (AmilGFP517a, b; clone 114c8 and EST C012-c9 clone) at a concentration of 0.35 μg/μl appeared orange-pink to the naked eye (Figure 4.5B, tube 1). The absorbance spectra showed two double component curves with peaks at 506 and 505 nm and shoulders at 477 and 478 nm for AmilGFP517a and AmilGFP517b, respectively (Figures 4.6A and 4.6D, open orange circles).

The excitation spectra of these two proteins were characterized by peaks at 518 and 517 nm and shoulders around 490 nm for the AmilGFP517a and AmilGFP517b proteins, respectively (Figures 4.6B and 4.6E, blue filled circles). Their emission spectra had a major peak at 517 nm and a shoulder at 550 nm (Figures 4.6C and 4.6F, green circles).
Figure 4.6 Spectral characterization of amilGFP517a and amilGFP517b proteins in solution (PBS, pH 7.2) at room temperature.

All spectra were normalized relative to the main peak. The values are expressed as units on the principal y axis. The secondary y axis represents the fourth derivative value of the dependent variable with respect to the energy of the photon and is represented as the red line overlapping the spectra. A) Absorbance spectrum of amilGFP517a, denoted by “●”. B) Excitation spectrum of amilGFP517a, denoted by “○”. C) Emission spectrum of amilGFP517a, denoted by “○”. D) Absorbance spectrum of amilGFP517b, denoted by “●”. E) Excitation spectrum of amilGFP517b, denoted by “○”. F) Emission spectrum of amilGFP517b, denoted by “○”. The numbers in the graphs represent the values of the peaks to the left in nanometres.
4.2.4.2 Spectral signature of *A. millepora* red FP (AmilRFP602)

The AmilRFP602 (EST C004h7, pQE clone vblg274c) recombinant protein (0.35 μg protein/μl) appeared red-violet to the naked eye (Figure 4.5B, tube 2). The bacterial colony, which normally is creamy white in colour, was carmine-red in the centre of the colony when expressing RFP (Figure 4.5C, panel 2). The absorbance spectrum of this protein (Figure 4.7A) spanned from 450 to 580 nm and was characterized by a main peak and a shoulder at 568 and 529 nm, respectively.

The excitation spectrum of the RFP (Figure 4.7B) spanned from 500 to 620 nm. The spectral curve was defined by two main components, a peak at 584 and a shoulder at 540 nm. The emission signal of this RFP spanned from 580 to 700 nm. The curve was defined by a single peak at 602 nm; however, results from derivate analyses showed heterogeneity under the curve and predicted up to three components, which explained peaks with maxima at 582, 591, and 616 nm, respectively.
Figure 4.7 Spectral characterization of recombinant AmiRFP602 in solution (PBS, pH 7.2) at room temperature.

All spectra were normalized to the main peak. The values are expressed as units on the principal y axis. The secondary y axis represents the fourth derivative value of the dependent variable with respect to the energy of the photon and is represented by the red line that overlaps the spectra. A) Absorbance spectrum denoted by "○" B) Excitation spectrum denoted by "■". C) Emission spectrum denoted by "●". The numbers in the graphs represent the values of the peaks to the left in nanometres.
4.2.4.3 Spectral signature of *A. millepora* CP (AmillCP597)

The AmillCP597 (clone 287-pQE, clone 2:2-pGMT) recombinant protein (0.35 μg/μl) appeared blue to the naked eye (Figure 4.5B, tube 3). When expressing this protein, the bacterial colony (Figure 4.5C, panel 3) had a deep blue centre, characteristic of the blue/white selection cloning-assay, which is based on the reaction of the β-lactosidase enzyme with the substrate X-gal.

The absorbance spectrum (Figure 4.8A) spanned from 460 to 650 nm, defined by a peak and a shoulder at 597 and 563 nm, respectively. Results from derivate analyses showed a small contribution at 460 nm, and the shoulder had two components at 555 and 563 nm. The main peak was heterogeneous in composition and was defined by two peaks at 602 and 610 nm.

4.2.4.4 Spectral signature of *A. millepora* CP (AmilCP601)

The AmilCP601 (clone 281, EST clone c015-a8) protein (0.35 μg/μl) appeared almost identical to its CP597 counterpart (Figure 4.5B, tube 4), and the bacterial colony also appeared deep blue (Figure 4.5C, panel 3).

The absorbance spectrum (Figure 4.8B) spanned from 460 to 650 nm, defined by a peak and a shoulder at 601 and 566 nm, respectively. Results from derivate analyses show a small contribution at 461 nm, the shoulder defined by two components at 555 and 566 nm. The main peak appeared to be heterogeneous in composition and had two spectral peaks at 605 and 612nm.
Figure 4.8. Spectral characterization of two recombinant CPs measured in solution (PBS, pH 7.2) at room temperature.
The spectra were normalized to the main peak. The values are expressed as units on the principal y axis. The secondary y axis represents the fourth derivate value of the dependent variable with respect to the energy of the photon and is represented by the red line that overlaps the spectra. A) Absorbance spectrum from the protein AmilCP597 (clone 2:2) denoted by “○”. B) Absorbance spectrum from the protein AmilCP601 (EST clone c015-a8) denoted by “○”. The numbers in the graphs represent the value of the peaks to the left in nanometres.
4.2.4.5 Spectral characterization of the recombinant proteins under extreme pH conditions

To analyse chromophoric solvatation under acidic or basic conditions, a series of absorbance scans over the entire visible regions was performed using the same amount of recombinant protein at pH values of 3–12. Similar to "untreated" sample spectral characterizations, the x axis represents the wavelength in nanometres, and the y axis represents the absorbance in relative units.

The AmillGFP517b protein (pQE clone 167m1, EST c012-c9) normal absorbance spectrum was characterized by a main peak at 506 nm and a shoulder at ~470 nm (Figure 4.9B, black rhombus). Under acidic conditions, the protein showed a similar spectral shape (Figure 4.9B, blue triangles); the shoulder was maintained and there was a slight shift of 3 nm in the main peak. However, a new peak appeared in the blue region at ~390 nm under acidic conditions. In alkaline conditions, the protein had an absorbance maximum at 447 nm, and this peak was bell shaped (Figure 4.9B red circles).

The absorbance of the AmillRFP602 protein (pQE vblg274, EST c004-h7) under normal conditions was defined by a main peak, shoulder, and small peak at 568, 529, and 460 nm, respectively (Figure 4.9C, black rhombus). Under acidic conditions, the protein (Figure 4.9C, blue triangles) had different spectral characteristics, and the absorbance maxima shifted toward the UV-blue region. The main peak occurred at 380 nm, and a UV-shifted shoulder and two minor components appeared at 354, 455, and 556 nm, respectively. Alkaline conditions (Figure 4.9C, red circles) also affected the spectral characteristics of the protein. The spectral peaks moved toward the blue region but to a lesser extent compared to under the acidic conditions. The spectrum exhibited a single bell-shaped curve with an absorbance maximum at 454 nm and a small peak at 556 nm.

The AmillCP601 protein (pQE clone 281, EST c015-a8) normal absorbance spectrum was characterized by a maximum peak at 601 nm and a shoulder and smaller component at 566 and 462 nm, respectively (Figure 4.9A, black rhombus). The spectral shape did not change under acidic conditions (Figure 4.9A, blue triangles); however, there were 4-nm blue shifts in the main peak, shoulder, and small peak. In contrast, the spectral signature of the protein under alkaline conditions showed a more dramatic blue shift. Basic conditions resulted in a shift in the absorbance maximum to 452 nm; the peak appeared to be a single component and was bell shaped (Figure 4.9A, red circles). A small peak was observed at 557 nm.
Figure 4.9 pH-dependent behaviour of the recombinant proteins in solution at room temperature.

The spectra were normalized to the main peak. The values are expressed as units on the principal y axis. In all spectra, ‘’ , ‘’ , and ‘’ represent absorbance data collected under normal (pH 7.2), acidic (pH 4.0), and alkaline (pH 12) conditions, respectively. A) AmilCP601; B) AmilGFP515b; C) AmilRFP601. The numbers in the graphs represent the value of the peaks to the left in nanometres.
4.3 Discussion

4.3.1 Amino acid substitutions and the functions and phenotypes of the protein

Scattered substitutions in the primary structure of the proteins were observed in amino acid alignments of similar clones inside each contig (Figures 4.1 and 4.2). The expressed clones also revealed substitutions that determined their fluorescent phenotype, e.g., the position H142N of the AmilCP601 and AmillCP597 proteins (Figure 4.3 and 4.10). The imidazole ring of histidine allows it to act as either a proton donor or acceptor at physiological pH. Hence, it is frequently found in the reactive centre of enzymes. This position is also variable in several other GFP-like proteins and, to some extent, is related to the phenotype of the protein. For example, in the hydrozoan GFP and several fluorescent anthozoan-related molecules, this position is occupied by histidine and serine, respectively (Prasher et al. 1992, Labas et al. 2002, Gurskaya et al. 2003, Tu et al. 2003, Karasawa et al. 2004, Sun et al. 2004, Luo et al. 2006), whereas in non-fluorescent CPs from hydrozoans and anthozoans, this position is mainly occupied by threonine, alanine, or cysteine (Gurskaya et al. 2001, Shagin et al. 2004). The importance of this position for generating fluorescent mutants from CPs cloned from sea anemones and corals was assessed by changing it to serine. The mutation generates a far-red (615–645 nm) FP that differentiates it from non-fluorescent ones (Lukyanov et al. 2000, Gurskaya et al. 2001).
Figure 4.10 Amino acid alignment from representative GFP homologues across Metazoa.

The background shading pattern goes from red to blue for the more similar to dissimilar residues, respectively. Along the alignment, several Acropora variable positions are depicted. Information related to the NCBI accession number and reference of each sequence is summarised in Table 4.
Figure 4.10 Continuation
The difference between the two Acropora CPs at this position may account for the observed red absorbance peak shift from 597 to 601 nm. Scleractineans (Figure 4.10) mainly have serine in this position, and only A. palmata shares the same residue as A. millepora, suggesting a deep genera-related substitution. Similar comparisons of cloned GFP-like molecules in other metazoans have shown no anthozoans that share similar amino acids, with the exception of a fluorescent hydrozoan homologue (Prasher et al. 1992, Shagin et al. 2004, Luo et al. 2006). It is tempting to speculate that this class-related amino acid residue is a key ancestral position retained in the majority of hydrozoans and is derived from anthozoans.

Aspargine 142 in the amilCP gene is widespread in Acroporids and is also a shared residue in CPs of Goniopora tenuidens, which belong to a suborder of Acropora. This implies that, similar to its anemone-like homologues, the position is characteristic of the scleractinean CP gene. However, all fluorescent-like proteins have Cys148, a substitution observed in one favid and several anemone fluorescent CPs (Gurskaya et al. 2001, Shkrob et al. 2005).

The E26V substitution in the amilCP protein is located near the 142 or 143 site; however, the substitution falls just below the equatorial region (Figure 4.3), and the 142,143 sites are above the equator region. The substitution appears to have electrostatic properties in that protein surface region, as suggested by the isoelectric points of E26 (3.2) and V26 (5.96). The proportion of proteins with the V26 residue was 0.12 compared to 0.88 for those with the E26 amino acid (Figure 4.1). Compared to the rest of the scleractinians (Figure 4.10, general alignment), the consensus amino acid is glutamate (Ando et al. 2002, Karasawa et al. 2003, Carter et al. 2004, Karasawa et al. 2004, Sun et al. 2004, Kogure et al. 2006). However, there are fewer FP-like sequences from the Montastraea group, and Scolymia cubensis and Meandrina meandrites also show a similar V/E substitution (Kelmanson & Matz 2003, Carter et al. 2004, Sun et al. 2004). Comparisons of CPs and FPs from Acropora show that threonine occupies this position in all FPs (NCBI submitted sequences, Table 4). A similar position was found in sequences from distantly related M. cavernosa and Favia favus corals, indicating an ancient origin/fixation (Sun et al. 2004, Tsutsui et al. 2005).

With regard to the functional consequences of the E26V substitution, no connection with the genesis of the chromophore is expected. This is because the side chain faces the exterior of the protein; however, it is well known that certain residues
located in the surface of anthozoan GFP-like proteins are responsible for oligomerization (Baird et al. 2000, Wall et al. 2000, Zacharias 2002, Pakhomov et al. 2004, Verkhusha & Lukyanov 2004). The substitution, when compared with data from native and mutated proteins, is not located in the “interface” zone and therefore is not likely to be involved in the formation of the tetrameric complex. Alternatively, this amino acid might function in the binding/interaction with other proteins. It is possible that this protein region may promote its targeting to specific subcellular locales.

S53A was another substitution observed in the two CP genes. The dominant position was found to be alanine, with a relative proportion of 0.88 (Figure 4.1). The substitution was located at the base of the α-helix; hence, a low electronic interaction with the chromophore is expected to occur (Figure 4.3). The isoelectric point change between substitutions is small (6 to 5.68). Compared with the rest of the scleractinians (Figure 4.10), this substitution was also commonly alanine, and serine was the only other amino acid found in the coral-derived proteins. However, serine was the consensus in the FP contig (Figure 4.2), indicating a possible origin of the substitution in the CP gene and a fixation in the FPs. Further comparisons with the rest of the metazoans revealed that nearly all anthozoans as well as bilateria-derived molecules possess this A/S variation; the hydrozoan jellyfish was the only group to have a proline at this position (see Figure 4.10 and Table 4 for references). It is likely that this position plays a relatively important structural or functional role because of its restricted variability. As mentioned previously, the amino acid may not be entirely related to the acquisition of the colour/fluorescence phenotype.

From non-expressed CP variants, the S206L substitution was also abundant in the contig (Figure 4.1). What makes this interesting is that this position is relatively close to E26V; it also faces the outside of the barrel. Furthermore, according to the oligomer interaction/formation map, position 206 is not involved in tetramer formation, suggesting that it provides a larger area for facilitating protein–protein interactions.

In relation to the variation in the FP contig, the R10K substitution observed facing the exterior of the protein was highly abundant (Figure 4.2), with lysine being more abundant in the contig. Compared with other scleractinians (Figure 4.10 general alignment), lysine is the predominant amino acid at this position. Arginine at this position was also abundant in the FP contig; however, no coral GFP-like protein was
observed to have this amino acid, making the substitution especially unique to the AmilGFPb protein. The isoelectric points of the amino acids differ by 1.56 units, and they are not involved in tetrameric protein formation. It is possible that these amino acids are targets for selective recognition by similar proteins. An analogous situation may occur with the A13T substitution; however, structural changes in the surface of the protein are expected with this type of substitution.

A more important substitution with respect to the colour/fluorescence phenotype in the FP contig is D66T (Figure 4.2). This position is in the tripeptide motif involved in the generation of the chromophore. The change in green and red fluorescence observed in the two expressed variants (Figures 4.6C and 4.7C) is explained by this substitution. Threonine at this position is very rare in scleractineans and is found only in *M. meandrites* GFP (Figure 4.10). It is also rare in the rest of the metazoans; Phalidium GFP is the only non-anthozoan FP bearing this amino acid (Shagin et al. 2004, Sun et al. 2004). Aspartic acid, in contrast, was poorly represented in scleractinian GFP-like proteins and was observed only in distantly related *Favia* and *Montastraea* corals (Carter et al. 2004, Tsutsui et al. 2005).

P166A was a widespread surface variation in the contig (Figure 4.2). Two expressed green fluorescent variants had proline (Figures 4.4A and 4.4B), whereas the red fluorescent variant had alanine (Figure 4.4C). Overlapping the coralimorpharian DsRed map with the Acropora FP map shows that this substitution falls in the polar monomer–monomer interface (Wall et al. 2000). It is clear by comparing the substitution among the scleractinians (Figure 4.10) that no other coral bears this substitution; however, more extensive comparisons with sea anemone-, hydrozoan- and bilaterian-derived GFPs have shown that proline is present in the OFP of the octocoral *Cerianthus*, in the GFPs of the sea anemones *Anemonia sulcata* and *Heteractis magnifica*, and in the GFP of the coralimorpharian *Ricordea florida* (Wiedenmann et al. 2000, Tu et al. 2003). It is interesting that alanine was found only in non-cnidarian GFPs of the copepods *Pontellina plumate* and *Chiridius poppei* (Shagin et al. 2004, Masuda et al. 2006). There are two possible explanations for the similarities of this position with distantly related taxa. First, the presence of such substitutions in Acropora FPs represents an ancestral variation that may be related to the state/formation of the tetramer complex. Second, these similarities could be interpreted as *de novo* creations rather than inherited information from early ancestors because the Acropora genus is relatively recent in origin.
Two consecutive substitutions, R180Q and S183T, occur externally in the polar interface and internally in proximity to the chromophore, respectively, of the Beta-can (Figure 4.4). The difference in isoelectric point for the external substitution is 5.85 units, whereas that for the internal substitution is 0.02 units. R180 was scarce among the scleractineans and was found only in the GFP from Acropora sp (Figure 4.10), whereas glutamine at this position is dominant in Acropora genus GFPs. Glutamine at this position is also present in the closely related Pocillopora damicornis CFP and in distantly related proteins from Fungia concinna (Carter et al. 2004, Karasawa et al. 2004).

The internal S183T substitution appears to be widely distributed among scleractinians. The distal hydroxyl residue of the serine side chain may have some effect in the electronic environment of the chromophore and may account for the observed red shift in fluorescence between the expressed proteins. Two interesting variants (EST clones c002-d8 and a011-h9; Figure 4.2) in the FP contig showed the typical green TYG chromophore; however, threonine at this position instead of serine resulted in a red shift similar to that observed in the RFP with the DYG chromophore and Thr183. Based on the literature, the AmillFP variation at 183 does not fall strictly into the amino acids that interact with the chromophore. Despite being internally buried, this amino acid is far from the supposed interacting position (Wiedenmann et al. 2004).

4.3.2 Acropora GFP-like protein electrophoretic signature

The theoretical molecular mass of the FPs and CPs were 27.6 and 26.3 kDa, respectively. The apparent masses were in agreement with these values when the proteins were heat-denatured in the presence of a reducing agent (Beta-mercaptoethanol; Figure 4.5A). Values found in the literature for several GFPs, including the former avGFP, Acropora formosa "pocilloporin" (abs max 580 nm), Pocillopora damicornis GFP-like protein (abs max 560 nm), and many other anthozoa- and bilateria-based proteins, similarly were in the ~28 kDa range (Ward & Cormier 1979, Dove et al. 1995, Niwa et al. 1996, Dove et al. 2001, Wiedenmann et al. 2002, Pakhomov et al. 2004, Masuda et al. 2006). There were, however, a few extra bands in similar contig-related proteins. The CPs (Figure 4.5A, lanes 4 and 5) had extra bands that appeared to be tetramers and dimers. It is possible that the
oligomeric interaction was stronger in the CPs than the FPs and that heat denaturation did not fully dissociate the protein complexes.

The FP proteins (Figure 4.5A, lanes 1–3) did not show high molecular mass bands after being subjected to heat denaturation treatment. This is consistent with the idea that FPs form less stable oligomers compared to CPs. The amilRFP protein displayed a band at 10 kDa, which probably represents a protein fragment; backbone fragmentation occurs near the chromophore as a consequence of its maturation (Wiedenmann et al. 2000, Martynov et al. 2003). Although these assumptions could be tested by examining the crystal structures of each protein, this is outside the scope of this study.

4.3.3 The AmilRFP602 spectral signature: similarities and hypotheses

The amilRFP602 protein in PBS possesses dual component excitation spectra with a peak and shoulder at 584 and 540 nm, respectively (Figure 4.7B). However, the incipient eluant containing this protein had twice as much NaCl than was recommended by the manufacturer. Although the double component excitation spectra were observed, the shoulder of the PBS-eluted protein was the main peak, and there was also a component that absorbed at 580 nm (Figure 6.1, Appendix). This double component signature has been observed in the A. victoria GFP. In hydrozoan proteins, these two components correspond to the protonated and deprotonated states of the chromophore (Miyawaki 2004). The different states of the chromophore possess different emission characteristics when the protein is illuminated with UV and blue lights; however, these states have also been linked with proton pumping behaviour (Agmon 2005). It is uncertain whether the amilRFP protein has the capability to pump protons under the ionic conditions used in these experiments. For this study, the comparison/description of the GFP-like proteins that eluted in PBS at pH 7.2 remained the subject of focus.

The larval excitation signal (measured in situ and from the cloned product) has differences in peak excitation maxima compared with its counterparts from adult tissue (Cox et al. 2007, Alieva et al. 2008); the larval protein exhibits a red shift of ~24 nm. The absorbance characteristics of the amilRFP602 protein were very similar to those observed in some non-fluorescent GFP-like proteins from Anemonia sulcata (Wiedenmann et al. 2000, Bulina et al. 2002), i.e., a double component curve and an
absorbance maximum of 560 nm. It is interesting that the adult version of the protein (amilRFP593) described by Cox et al. (2007) and Alieva et al. (2008) had an emission signal that was blue-shifted by 9 nm compared with the larval protein described here. Whether this spectral difference confers special attributes to adult metabolism or coral development remains unclear.

Lastly, similar to the excitation signal, the emission signal of the larval protein under stronger ionic conditions was shifted by 8 nm (Figure 6.1, Appendix). Regardless of the ionic conditions at the time of measurement, it is clear that the emission spectrum of this RFP shows the largest red shift of its class (Figure 4.7C). The actinarian-derived RFP from *E. quadicolor* showed a similar far-red emission at 611 nm, and a mutant product from one hydrozoan CP called KillerRed also had a similar emission maximum (610 nm; Wiedenmann et al. 2002, Bulina et al. 2006).

**4.3.4 Larval AmilGFP517a, b: The most common emission for a coral GFP-like protein**

It is interesting that an emission maxima of ~517 nm (Figure 4.6C, F) appears to be one of the most abundant spectral signatures detected in caribbean corals (Fux & Mazel 1999). However, spectral data from adult *A. millepora* colonies showed the presence of two green variants with emission peaks at 504 and 512 nm (Cox et al. 2007). These small spectral differences between the larval and adult proteins appear to be based in the amino acid composition (Alieva et al. 2008; Figure 4.10). In a phylogenetic analyses, the adult GFP protein clustered with the cyan FP rather than the larval GFP (Figure 3.2a, chapter 3). These data suggest that some spectral channels are under selective pressure.

The excitation peaks of the adult green FPs were blue-shifted to ~504 nm, whereas the peaks of the larval FPs were at ~517 nm. Using the fourth derivate of the spectral data, it was possible to detect components of the excitation spectra that were similar to those of the green-yellow mutant zFP538-D68N (Ianushevich lu et al. 2002) cloned from a sea anemone (*Zoanthus* sp.). Before publication of the adult-derived *A. millepora* FPs in the NCBI gene bank, the anemone FP genes were closest at the amino acid level to the Acropora protein, which explains the similarities in their spectral signatures.
4.3.5 The amilCP products: similarities with other cnidarian proteins and potential FPs

Spectrally, AmilCP597 and AmilCP601 (Figure 4.8A, B) appear very similar to the aeCP597 protein (Shkrob et al. 2005) cloned from the blue margins of the pedal disc of the sea anemone *Actinia equina*. In *A. millepora* adult colonies, this colouration is distributed principally in the growing tips of the coral colony. In planula larvae, the blue colouration due to the presence of these proteins is distributed across the endodermal tissue. In a recent attempt to generate new biotechnological probes, a mutant gene of aeCP597 containing a few amino acid substitutions exhibited far-red fluorescence, which may be useful for multi-labelling purposes because of its non-overlapping emission characteristics. Similarly, the *A. millepora* CP genes could be manipulated by mutagenesis to express a red-shifted FP.

The creation of a far-red mutant from Acropora could be advantageous if a stable coral cell line becomes available. During spawning events, it may be possible to deliver the far-red mutant gene into the early dividing embryo cells. Difficulties in translation would be avoided because the probe is part of the same transcriptome and would use the same codons.

The photoinduction of fluorescence was investigated using recombinant CPs. As shown in Figure 4.5D, panel 1, similar amounts of protein were blotted onto 3M paper and then left under a 100 W Hg UV, blue, or green light power beam for 5 min. Proteins not exposed to light served as controls. No light effects in the red or yellow-green fluorescent state were observed (Figure 4.5D, panels 2 and 3). Once adsorbed by the paper, the solvent evaporated, leaving the proteins in a dry state. The dried proteins acquired fluorescence, which suggests that water aided in quenching the fluorescence of the solvated proteins.

Consistent with this, in *A. victoria* GFP, a complex network of water-mediated hydrogen bonding occurs between the chromophore and its surroundings; in particular, the hydroxyl groups of Try66 and Thr203 participate in hydrogen bonding (Ormo et al. 1996, Tsien 1998, Follenius-Wund et al. 2003). The elimination of the water may affect the internal charge environment of the protein. This level of dehydration is difficult to achieve in an aqueous cell environment such as the cytoplasm. However, if the protein is secreted to a hydrophobic environment (such as lipid vesicles), it is possible that these CPs may then fluoresce. *A. victoria* GFP is
expressed mainly in the endodermal planulae, and this tissue generally has a high lipid content that is used for energy. The endodermal fluorescence could be explained in part by the presence of CP-like proteins that were dehydrated and subsequently acquire a fluorescent state. It is important to note that FPs in the green state undergo one dehydrogenation reaction and that two dehydrogenation reactions occur with proteins in the red state during chromophore coordination. An extended dehydrogenation reaction may have occurred with the blotted CPs; this might have been aided by the physical removal of water and/or oxygen-mediated catalysis, which are unlikely to occur in nature.

4.3.6 Chemical nature of the GFP-like chromophore: hypsochromic pH-dependent shift

The three main GFP-like recombinant variants were subjected to one-step pH changes. The results showed similar (473–446 nm) blue shifts that occurred under alkaline conditions in all proteins; however, under acidic conditions only the RFP variant showed a greater shift toward the UV region (Figure 4.9). These extreme pH conditions cause denaturation of the protein, which exposes the chromophore directly to the solvent. Through its spectral characteristics, the type of chromophore can then be determined (Martynov et al. 2003, Pakhomov et al. 2004). The absorbing behaviour that occurred with pH change for the AmilGFP517 protein appears to be very similar to data from a Pectinidae coral protein called Dronpa (Ando et al. 2002, Ando et al. 2004). Furthermore, the normal absorbance at pH 7.4 appears to overlap with a peak at 503 nm. In acidic conditions, Dronpa displays a peak at 390 nm with a proportional decrease in the major (503 nm) peak, and this is also observed in the Acropora protein (Figure 4.9B). The 390 and 503 nm peaks correspond to the neutral and ionised states of the phenolic hydroxyl group of the chromophore (Tsien 1998).

The Acropora CP and RFP show a very similar alkaline-related peak (Figures 9A and 9C). A comparable behaviour has been recorded for Favid corals and implies that the chromophores share very similar chemical structures (Oswald et al. 2007). Detailed analyses of chromophore biogenesis have confirmed that some red FPs and non-fluorescent CPs share similar pathways in the three-step reaction that is responsible for the formation of the red-shifted chromophore (Gurskaya et al. 2000, Kelmanson & Matz 2003, Martynov et al. 2003, Shkrob et al. 2005).
4.3.7 The 3D structure of GFP and its nearest structural protein neighbour

The topological variability of GFP-like proteins is important for supporting the nervous systems of many distantly related organisms. Hopf et al. (2001) explored the topological similarity of perlecan binding of nidogen 1 and the GFP Beta-barrel shape. Despite the low amino acid similarity (10%) in the normal alignment, the overall shapes of both proteins appeared to overlap. Nidogens are a component of the basement membrane, which is a thin sheet made of extracellular matrix that separates tissues such as smooth muscle, skeletal muscle, parineural cells, lymph nodes, and fibroblasts. The building of such membranes involves other proteins like collagen IV, laminins, and proteoglycans. Collagen IV and laminins self-assemble in a network-like style, whereas the nidogens connect these networks. It would be interesting to compare the expression patterns of Acropora nidogen-like proteins and GFP-like proteins. The similarities in the β-barrel domain may also explain other mutations not related to the colour/fluorescence phenotype. The spatial distribution of some amino acids (Figures 4.3 and 4.4) on the surface of Acropora GFP-like proteins suggests the presence of an electrically charged “patch” environment. The surface anomalies of the clones and the binding-station phenotypes are similar to the nidogen proteins, and it may be possible that they can bind functionally related proteins.

4.4 Conclusions

A total of five recombinant proteins—two green FPs, one red FP, and two CPs—were characterised and compared with data from other related proteins. Analyses of the GFP-like protein diversity from the Acropora EST database showed important and abundant variants that likely contribute to the colour diversity that occurs in the embryo stage of *A. millepora*. The molecular masses of FPs were slightly larger than those of CPs but were in the range of other cloned GFPs from a variety of taxa. The spectral characteristics of the recombinant proteins showed a typical double excitation peak for the RFP and a widespread taxa-related green emission with unusual excitation wavelengths. Absorbance of the CPs was similar to other red-shifted cnidarian proteins, and, when blotted onto 3M paper, fluorescence occurred.
CHAPTER 5

TEMPORAL AND SPATIAL mRNA EXPRESSION AND FLUORESCENCE ANALYSES OF GFP-LIKE GENES DURING ACROPORA MILLEPORA EMBRYOGENESIS

5.1 Introduction
During metazoan embryogenesis, several genes that control morphogenesis appear to share a common spatial pattern of expression. It is during this time that organs are formed, but apart from organisms that undergo metamorphosis, morphogenesis occurs only once during the lifetime of an organism. For organisms that do undergo metamorphosis, a secondary morphogenesis event occurs. During this time, in some hydrozoans, the cellular fate of the planula ectoderm is changed by apoptotic mechanisms, creating new tissue characteristics for the newly created “polyp” (Seipp et al. 2001). This polyp differs from its planula stage in terms of its sessile character and, for some reef-dwelling cnidarians, the polyp stage is the starting point of exoskeleton formation.

Due to advances in mRNA manipulation, now it is possible to visualise the in situ expression of a specific gene in model organisms. This technology, when applied during embryological development, allows direct comparison of dynamic changes in gene expression across different animal phyla. With such information, the interpretation of possible evolutionary pathways becomes possible. In the coral Acropora millepora, for example, the expression pattern of important developmental genes, such Snail (involved in metazoan early gastrulation) and Dpp (involved in dorsoventral specification), are similar to those observed in the bilaterians Drosophila and mice (Ball et al. 2004; Hayward et al. 2004). Although cnidarians have typically been categorised as radial rather than bilateral animals, these expression patterns suggest that the common cnidarian-bilaterian ancestor was already a bilateral organism.

Genes involved in mesodermal tissue specification have also been found in diploblastic cnidarian genomes. The existence of such genes in a diploblastic
organism, which in higher metazoans specify the third tissue layer, has been explained in terms of potential initial/ancestral cellular (proliferation) activity that was later recruited/exploited in a triploblastic bauplan to create a new type of tissue (Technau 2001; Technau & Scholz 2003).

Symbiotic corals have several reproductive strategies, products of the combined effects of life history traits (Hall & Hughes 1996). They can reproduce asexually (via fragmentation) or sexually, in which hermaphroditic or gonochoric species fertilise gametes and development occurs internally (brooders) or in the water column (broad spawners) during synchronous mass spawning (Harrison et al. 1985). During gamete release into the water column, potential hybridisation between different species may have given rise to the diversity observed today, especially within the genus Acropora (van Oppen et al. 2000, 2001, 2002).

A. millepora is one of several coral broad spawners in the Great Barrier Reef. This species, along with others, usually spawns during the second or third night after the spring full moon, at around 9–10 p.m. Fertilisation occurs at the water surface and the first cell divisions begin 1–2 h later. After 12 h, the embryo remains at the surface and by then its morphology resembles that of an oriental prawn chip, an irregular milky brown-coloured bi-layered array of cells about 200 µm in diameter. Occasionally, it is possible to observe a red-pink colouration at this stage in the embryos (Figure 5.1; 12-h prawn chip embryo).
Figure 5.1 Embryological development of *Acropora millepora*. Modified from Ball et al. (2002).
After approximately 24h post-fertilisation, the embryo usually remains non-motile, and its general morphology is similar to that of a fat donut, a product of the gastrulation-like process that forms the endodermal tissue and the oral pore (Figure 5.1, 22-h “fat donut” embryo). It is during the second night that gastrulation usually ends and the embryo acquires a spherical shape and rotation behaviour, driven by synchronous ciliary beating, begins.

From approximately 36 h post-fertilisation to the next day, no dramatic morphological changes are observed, but there is a slight elongation along the axis, forming a pear-like shape (Figure 5.1, 72 h post-fertilisation). Swimming behaviour is now directional (spinning on its own axis), with the aboral end first; however, the embryos are still close to the water surface, because of their high lipid content.

From this point to approximately 88–96 h post-fertilisation, the embryo experiences a dramatic elongation along its oral–aboral axis, acquiring the classic planuloid shape, with the appearance of diverse cellular types, including neurons, nematoblasts (nematocysts), and specialised secretory cells. In addition, thickening of the aboral-most ectodermal tissue is visible (Figure 5.1, 96-h planula larvae; Ball et al. 2001, 2002). At this point, the coral larvae generally leave the water column, actively following chemical cues leading them to the bottom of the reef. This bottom-searching period can be variable, but usually lasts approximately 2 days (Dr. Andrew Baird, JCU comp pers). After the planula finds the right spot to settle, it will attach via its aboral end, initiate metamorphosis to form a polyp, establish symbiosis, and begin exoskeleton formation (Figure 5.1, top view of two polyps). From then on, the coral polyp divides asexually until it reaches maturity (approximately 3–5 years). Then, it waits for the proper environmental cues to trigger the expulsion of its gametes into the water column, to continue the survival of the species.

This chapter describes the behaviour of green fluorescent protein (GFP)-like proteins in situ through the analyses of mRNA and protein expression at various stages during A. millepora development. Due to the ubiquitous fluorescence characteristics of the protein products, this combined method offers an excellent opportunity to observe the dynamics of gene and protein expression during the formation of a new individual. The information derived is discussed in terms of differences and similarities (observed in the pattern of expression and the cell type involved) with other gene expression patterns observed in this species and other cnidarians,
principally during the planula stage. Despite the great morphological/ecological plasticity of adults in this phylum (e.g., jellyfishes vs. corals), the planula bauplan is similar across the phylum, allowing proper comparison. Using this temporal and spatial information, in conjunction with other evidence, the potential roles of specific genes and proteins will be explored at the organism and cellular levels.

5.1.1 Statement of Goals
Expression patterns in space and time can provide critical information about potential functions of genes during embryonic development, and studies of this kind are particularly important in non-model organisms where traditional genetics approaches cannot be applied. Based on expression data it is possible in principle to assign the relevance (timing and level of transcription) and putative domains of functionality (e.g. if the transcript or protein is localized at the aboral end and in a distinctive neuron-like cells, it is possible that the domains of functionality could be related with signal transduction) of the gene or protein. Gene expression analyses are commonly performed using in situ hybridization techniques, however, thanks to the inherent fluorescent properties of the gene products, GFP-like protein expression can be analysed in vivo by the use of fluorescence microscopy. In this thesis, the combination of these two approaches was extremely useful due to the potential for cross hybridization between mRNAs encoding spectroscopically distinct proteins.

5.2 Results
5.2.1 The major GFP gene types are differentially expressed in early coral development

The expression of the fluorescent protein (FP) and chromoprotein (CP) loci was investigated during embryonic and larval development in Acropora by Northern blot analyses and whole-mount in situ hybridisation. High levels of nucleotide sequence identity within the CP and FP clades effectively precluded resolution between specific spectral variants (red fluorescent protein [RFP] vs. GFP). Thus, the mRNA expression patterns reported here reflect general characteristics of the CP and FP classes.

Northern blot analyses showed major differences in the timing of CP vs. FP expression (Figure 5.2H). While CP mRNA expression was detected from the
gastrulation (35 h) stage, mRNAs corresponding to the FP types were only detectable much later in development (the pear stage at 52 h; Figure 5.3H). *In situ* hybridisation experiments indicated that CP expression was predominantly, but perhaps not exclusively, endodermal during larval development. Expression was first detected during gastrulation, in developing endodermal tissue and scattered ectodermal cells (Figure 5.2, panels A and B, respectively). All positive cells in the endoderm and the scattered ectodermal pattern remained from this stage to the planula stage (Figure 5.2, panels C–E). After metamorphosis, strong CP expression was observed throughout the endoderm, but was most prominently associated with the developing mesentery walls (Figure 5.2G). In older polyps, CP expression was specifically associated with the endoderm in the region of the developing tentacles (Figure 5.2F).

*FP* probes, on the other hand, showed a highly specific pattern of expression in the ectoderm during larval development. At the pear stage, staining was associated with a subset of transectodermal cells that appeared to be denser in the aboral half of the larva (Figure 5.2, panel I). In early planulae (~52 h post-fertilisation), the ectoderm tissue that formed the mouth started to stain (Figure 5.2, panel J, arrowheads); slightly later, there were two obvious domains of expression, the extreme oral ectoderm inside the mouth and the dense transectodermal staining at the aboral end of the planula (Figure 5.2, panel K). After metamorphosis, the aboral staining disappeared, whereas the oral staining remained; strong endodermal staining was also observed in the majority of surveyed organisms (Figure 5.2, panel L). In older polyps, the endodermal signal was primarily associated with the forming mesenteries (Figure 5.2M).
Figure 5.2 In situ and in vitro time-dependent detection of the GFP-like mRNAs in Acropora development.

A, B) A. millepora chromoprotein (AMILCP) mRNA detected in a cross section and the whole 35-h embryo, respectively. C, D) Pear-stage AMILCP-associated detection observed under normal light and fluorescent enhanced view of the ectodermal surface, respectively. E) AMILCP was detected in the planula (approx. 88 h). F, G) Early and later expression of AMILCP in metamorphosed polyps, respectively. H) Northern detection of the AMILCP and fluorescent protein (AMILFP) genes in Acropora development. I, J) Early and late detection of the AMILFP mRNA in pear-like embryos. K) Detection of the AMILFP in planula larvae. L, M) Early and later staining pattern of AMILFP in young polyps. Scale bar equals 200 μm (Beltran-Ramirez et al. 2009)
5.2.2 Fluorescence microscopy of GFP variants

Although \textit{in situ} hybridisation was unable to resolve GFP and RFP mRNA expression, the strong and specific spectroscopic characteristics of the corresponding proteins enabled their distribution patterns to be determined via fluorescence microscopy. To supplement the \textit{in situ} hybridisation data, we examined the fluorescence characteristics of coral larvae during development.

After 52 h of development, at the pear stage, the majority of the embryos started to show a red colouration at the aboral end (Figure 5.3, panel A). Green fluorescence was detected throughout the endodermal cavity and was restricted ectodermally to the mouth area (Figure 5.3, panel D). At this stage, embryos observed under a CY5 filter showed variable low endodermal red fluorescence; ectodermally, all embryos showed strong cell-specific red signals towards the aboral end (Figure 5.3, panel G). Using a combined filter, embryos showed yellow endodermal fluorescence, because of the overlapping green and red signals, whereas tissue forming the oral and aboral ends appeared green and red, respectively (Figure 5.3, panel J).

The strongest GFP-like mRNA signal was detected in the planula stage via Northern analyses; \textit{in vivo}, larvae displayed a blue endodermal colouration, due perhaps to CP expression, whereas the red ectodermal aboral colouration observed previously was more evident (Figure 5.3, panel B). Planulae observed under GFP2/FITC filter showed strong green fluorescence throughout the endoderm, as well as within the ectodermal mouth and the surrounding area (Figure 5.3, panel E). Detailed inspection of the oral and mid-ectodermal areas showed tightly packed columnar cells forming the mouth sphincter, and isolated neuron-like cells running along the ectodermal basement, both with associated green fluorescence (Figure 5.3, panels M and N, respectively). A red fluorescent signal was detected in specific ectodermal cells with higher abundance towards the aboral end, forming a gradient along the planula body (Figure 5.3, panel H). Closer observation of the mid-ectodermal section under the CY5 filter showed columnar-type cells, with a broad basement layer, with observable basal prolongation and small spherical processes facing the larvae surface (Figure 5.3, panel O).
Planulae observed using the combined filter typically showed yellow endodermal fluorescence, a product of overlapping green and red signals (Figure 5.3, panel K). At greater magnification, the ectodermal mid-body section showed a greater abundance of RFP-expressing cells and isolated GFP-expressing ones. These cells also exhibited centrally positioned nuclei and columnar morphology with spherical processes facing the surface of the planula, and they appeared to extend along the ectodermal basement (Figure 5.3, panel P).

After metamorphosis, the polyps appeared translucent, with zooxanthella in their endodermal tissue (Figure 5.3, panel C). In young polyps observed under the GFP2 filter, green fluorescence appeared to be predominantly associated with the developing mesenteries, whereas it was localised to the mouth area in the ectoderm (Figure 5.3, panel F). Similarly, newly metamorphosed polyps observed under the CY5 filter showed primarily red fluorescence associated with the mesenteries, but no cell-specific ectodermal signal was observed (Figure 5.3, panel I). Later, after the polyps started to form the exoskeleton, green/red endodermal fluorescence was observed; however, the green mouth ectodermal signal remained in the majority of the polyps (Figure 5.3, panel L).
Figure 5.3 *In vivo* observation of the principal morphological and fluorescence characteristics of *Acropora millepora* embryogenesis.

A–C) Pear, planula, and polyp observed under normal lighting conditions, respectively. D–F) Pear, planula, and polyp observed under the GFP2 filter, respectively. G–I) Pear, planula, and polyp observed under the Cy5 filter, respectively. J–L) Pear, late planula, and polyps observed under the combined (WIB) filter, respectively. M) Detail of planula’s mouth, observed under the GFP2 filter. N–P) Mid-section of planula ectoderm, observed under the GFP2, Cy5, and WIV filters, respectively. Oral opening is oriented to the left in the larval pictures. Arrowheads indicate important features, such as specific cells or domains with associated fluorescence. Scale bars at the organism (A–L) and cellular (M–P) levels indicate 200 and 20 \( \mu \text{m} \), respectively (Beltran-Ramirez et al. 2009).
5.4 Discussion

5.4.1 GFP-like message temporal regulation

The “virtual Northern” has become an important comparative tool for relative gene expression comparisons during coral development (Anke et al. 2001; Hayward et al. 2001, 2002; Hislop 2003). The assay results demonstrate significant up-regulation of both Acropora GFP homologues during the planula stage (Figure 5.2H). In this planula stage, the organism is active and planktonic.

Comparing the levels of expression between polyp and planula (Figure 5.2, panel H), it is clear that GFP down-regulation occurs from the planula to polyp stage. Interestingly, in situ hybridisation revealed a considerable reduction in endodermal staining as the polyp became older and started to form mesenteries (Figs. 2 and 3, panels K and O, respectively), whereas the ectoderm seemed to lose its strong aboral cell-specific signal (Figure 5.2, panel L). In relation to the observed down-regulation of GFP-like proteins after metamorphosis, work using hydra as a model (Seipp et al. 2001) showed that almost all of the ectoderm of the metamorphosis-induced planulae underwent apoptosis, during which cell death and phagocytosis appeared to occur from the posterior to the anterior part of the planulae. These results indicated that the anterior cells that once functioned as a sensory organ on the active planulae had disappeared, to become the calicoblastic ectoderm, responsible for the deposition of exoskeleton. Similar to the case in Hydra, the observed changes in GFP-like transcript levels after Acropora metamorphosis may also be related to tissue characteristic modifications, from columnar cells in the planula to squamous tissue in the early calcifying polyp (Vandermeulen 1974, 1975).

5.4.2 Nature of the observed colouration in the early stages

According to the mRNA norther analyses, the GFP-like message appears after gastrulation (Figure 5.2, panel H). There were however, some cases when a colouration was observed in the proliferating (timing before the GFP mRNA detection) tissue of some embryos (Figure 6.3, Appendix). This coloration was attributed, instead of GFP-like based ones, to the presence of another type of pigment of maternal origin. The orange-pink pigment is likely a carotenoid compound. Carotenenes are small, soluble, organic, non-fluorescent molecules (Fingerman, 1965). The distribution of the pigmentation in the prawn chip embryo
was apparently on the surface of the cell, indicating its hydrophobic character. Also, during the 2006 spawning season, some of the eggs showed purple and orange pigments, which turned out to be insoluble in water, but soluble in ethanol (Wiedenmann Ulm Univ., personal communication).

5.4.3 Green-red signal distribution in planulae

Atoda (1951) reported on pigmentation in planulae from *Acropora bruggemanni*, in which the distribution of colouration appeared to follow a pattern similar to that in *A. millepora*: green at the oral end and red at the aboral end. More recently, a study by Hollingsworth et al. (2006) on the distribution and appearance of fluorescence in *Fungia scutaria* embryos described similar patterns to those observed in the planulae of *A. millepora*. In this previous study, the authors performed fluorescence microscopy and immunohistochemical analyses of protein extracts at specific time intervals (using an anti-hydrozoan GFP antibody). In general, they found up-regulation of these proteins occurring later in development. These results, compared with those of the virtual Northern (Figure 5.2, panel H) and fluorescence microscopy at later stages (Figure 3, panels J and K), are consistent with the purposed similarities in the up-regulation of mRNA expression and protein fluorescence relative to the age of the *A. millepora* embryo.

The mushroom coral planula, in comparison with *A. millepora*, does not display red transectodermal cells towards the aboral end (Figure 5.3, panels G and H), but it does display green fluorescence in the oral region (Figure 5.3, panels D and E). In *Fungia* embryos, the up-regulation of this oral green fluorescence was stronger as the planulae aged. *Acropora* green fluorescence in the mouth region was also similar, with increasing fluorescence at later stages. It has been hypothesised that green fluorescence in the oral tissue of the planula may function as an attractor for the dinoflagellate symbiont (Hollingsworth et al. 2005). This suggestion is based on indirect evidence of green phototaxis in free living coral symbionts. This phototactic behaviour has been also observed in other dinoflagellate species, each with its own spectral preference (Horiguchi et al. 1999). If such light preferences also occur in *Symbiodinium* clades, this behaviour may be correlated with slightly different mouth fluorescence signals. This hypothesis, however, may only be applicable in open systems where the symbionts are acquired from the environment; in closed systems
(i.e., maternally inherited), the presence of this signal may have other roles than the attraction of suitable algal strains.

Hollingsworth et al., (2006) also reported red fluorescence in *F. scutaria*, detected only in the endodermal cavity, and the up-regulation of this signal was less dramatic relative to its green counterpart. In *A. millepora*, the red fluorescent signal was also distributed in endodermal tissue, and this signal was up-regulated more quickly than in *F. scutaria* planulae. However, in both species, green fluorescence appeared first, and red fluorescence appeared later.

### 5.4.4 Non-endodermal *in situ* staining; no fluorescent product was observed

During our survey of fluorescence characteristics, some embryos exhibited variable endodermal signals, even though mRNA levels showed very little variation in the endoderm. Despite the correlation between mRNA and protein signals, these cases of under-detection of both signals could be explained by a combination of colourless protein production and undetectable fluorescence, and, perhaps, the down-regulation of the gene product (mRNA) with the acquisition of normal levels of protein per cell. To address the question, it will be necessary to systematically separate the variation by controlled coral crosses, and increase the sampling rate (fluorescent visual inspections on the developing embryos), particularly at later stages when the “normal” up-regulation would be expected to occur. To preferentially avoid interference by the ectodermal signal, small endodermal sections could be used for selective RNA detection/quantitation, relative to the actual state of tissue fluorescence. By doing so, it may be possible to determine whether the observed changes are actually a product of up-regulation of a colourless GFP-like protein, or due to the down-regulation of a specific fluorescent gene.

### 5.4.5 GFP-like protein-expressing cell types in *Acropora*

The first cell-specific signal occurred just after the end of gastrulation (early pear embryo, Figure 6.4, Appendix), but are more conspicuous at the pear stage (Figure 5.3 panel D, G). After this stage, diverse cell types became apparent (Ball et al. 2001, 2002, 2004). It appeared that the relatively low abundance of the newly appearing green transectodermal cells remained constant, based on closer examination of middle planula ectodermal tissue (Figure 5.3, panel P), whereas the
proliferation of red fluorescent cells began along with the elongation process (Figure 5.3, panel G) and became more pronounced towards the aboral end as the embryo became more elongated.

Closer examination of the green fluorescent cells distributed along the basement of the ectodermal tissue (Figure 5.3, panel N) showed strong nuclear signals and neuron-like cell morphology, based on long bipolar projections. Red fluorescent cells, on the other hand, were apparently absent from the basement of the ectoderm; instead, there were columnar-type cells, with small spherical processes at the edge of the planula surface, with a broad basement showing slender projections running along the ectoderm/mesoglea boundary (Figure 5.3, panels O and P). Both types seemed to be sensory neurons that reacted against RFamide antisera (Martin 1992).

In cnidarians, two general types of neurons can be found in the ectoderm: sensory cells distributed from the free planular surface to the mesoglea, characterised by the presence of an apical cilium and dense core vesicles, and ectoderm-basal ganglionic cells, distributed along the entire anterior–posterior axis of the planula (Kolberg & Martin 1988). These neurosensory cells are known to arise via the transdifferentiation of epithelia/muscle cells, whereas the ganglionic cells arise from migratory undifferentiated interstitial cells (Martin 1992).

GFP expression has been also linked with bioluminescence. In relation with this, photocyte characterisation in pennatulids (*Venterillum, Renilla, Stylatula, and Acanthoptilum*) and hydroids (*Obelia, Campanularia, and Clytia*) has revealed the presence of one or more cytoplasmic projections larger than about 20 μm. These ectodermal photocytes, according to the description, were also similar in shape to the *A. millepora* planula green fluorescent transectodermal cells (Figure 5, panel D). Based on this similarity and the relatively low cell-type diversity within the phylum (Miller & Ball 2000; Ball et al. 2001, 2002, 2004; Technau et al. 2005; Kamm et al. 2006), metabolic and functional similarities may also be expected. In the case of bioluminescent cnidarians, sensory neurons represent the cell type most capable of combining reception and control of a response via calcium homeostasis. Among the Scleractinian corals and their planula larva, however, no bioluminescence is observed. However, the presence of GFP-expressing cells with similarities to those observed in bioluminescent organisms suggests that although aequorin-like activity
was lost, GFP-like protein expression persisted throughout coral-specific evolutionary history.

Early detailed anatomical descriptions of *Acropora* planulae (Ball et al. 2001) showed that the mouth tissue, in contrast to the rest of the ectoderm, consisted of only one cell type. These distinct cells were of a slender, secretory type, with many vesicles along the cell body and ciliary processes at their ends, which presumably function in the creation of water currents to draw in small food particles. In the present study, most of the planulae observed under the GFP2 filter and stained with the amilFP probe *in situ*, showed strong signals associated with this tissue (Figs 5.2 and 5.3, panels K and M, respectively). A recent study in eukaryotic cells pointed out the specific compartmentalisation of GFP-like recombinant proteins within lysosomes (Katayama et al. 2008). Lysosomes are acidic compartments primarily dedicated to molecular degradation. That these secretory-like cells were fluorescent, combined with eukaryotic evidence of lysosome GFP-like protein compartmentalisation, suggests a role for amilGFP517 in digestive functions.

Closer observation of the red fluorescent cells (Figure 5.3, panel O) shows that, in addition to having neuron-like morphology, the cells contain several intracellular granules (Figure 6.5, Appendix) that were apparently highly fluorescent. Unfortunately, no data were available for comparison, at least at the planular level. However, several studies in adult corals, including some in *A. millepora*, have also reported these aggregations in actively GFP-expressing cells, and attributed this observation to the increased tendency of this protein to form oligomers (Salih et al. 1998, 2000, 2004; Cox & Salih 2005; Cox et al. 2007).

The endoderm also showed some fluorescence (Figure 5.3, panel N). The cells, however, were not easily recognisable as independent units; instead, a field of evenly distributed round, non-fluorescent bodies embedded within a relatively homogenous fluorescent region were observed, with occasional highly fluorescent, but small, spherical granulations. Combining the fluorescence and *in situ* data (Figure 6.6, Appendix) it is probable that the small granulations observed via fluorescence microscopy are nuclei, as the *in situ* hybridisation data identified similar small stained units. Few data are available regarding cell characterisation in the endoderm of the planula; instead, the cells are typically categorised as yolk cells, interstitial cells, nematoblasts, or gastrodermal cells capable of acquiring intracellular dinoflagellates that digest/distribute the food content (Spurlock & Cormier 1975; Ball
et al. 2002; Dandar-Roh et al. 2004). However, characterisation of G protein-coupled receptors (GPCRs) showed that these genes were expressed exclusively in the endodermal cavity, in a subset of cells described as round or ovoid-shaped and clearly different from yolk cells (Anctil et al. 2007). Thus, the GFP-like mRNA-expressing cells observed in this study may be the same cells expressing GPCRs. It would be interesting to perform a double in situ study to clarify this and to better characterise endodermal cell diversity.

Finally, regarding the green fluorescence associated with the ectodermal neuron-like cells, it is possible that this signal is not only due to GFP-like proteins. Considerable evidence has indicated that some neurotransmitters have fluorescent properties, including catecholamines and dopamine (Reynolds et al. 1981; Mabuchi Michiaki et al. 2001). The pure compound can readily be excited by ultraviolet–blue light; however, its detection in live material requires concomitant glyoxylic acid treatment (Battenberg & Bloom 1975; Kolberg & Martin 1988). Although the planulae were not treated in this way, the natural fluorescent properties of the neurotransmitters may contribute to the whole cellular signal. However, this study also confirmed the presence of true GFP-derived fluorescence, based on the similarities between the protein and mRNA signals detected in living and fixed cells.

5.4.6 Possible interaction of position-governing genes and differential GFP-like protein expression along the anteroposterior axis in Acropora embryos

Despite the great similarity observed in the phylogenetic comparison (Chapter 4), green and red fluorescent proteins were differentially expressed along the planula body (Figure 5.3, panels J and K). In situ amilFP detection was insufficient to resolve the mouth- or aboral end-specific expression; instead, both patterns were observed at the same time regardless of the probe used (GFP or RFP). However, detection of both showed that the signal was distributed at the extremities of the principal axis. Using this spatial information and previously published work related to the in situ detection of important developmental genes in other cnidarian planulae, it is possible that the GFP-like genes are being regulated by such genes.

For example, expression of Snail-Am in early gastrulation (Hayward et al. 2004) resembles both GFP-like expression patterns, in the sense that both appeared in the proliferating tissue that subsequently formed the endoderm. In other organisms, Snail
plays important roles in mesoderm specification during embryogenesis (Technau & Scholz 2003). In *Nematostella vectensis* and *Hydra* embryogenesis, *Wnt* and *Brachyury* expression patterns show some similarities with *Acropora amilFP* expression, specifically in the oral pore domain (Technau 2001; Kusserow et al. 2005). These transcription factors are involved in mesoderm specification and the blastoporal signalling centre (Technau & Scholz 2003). Hydrozoan *Hox*-like gene expression overlaps with aboral RFP expression in *Acropora* (Masuda-Nakagawa et al. 2000). *Hox* genes are primarily responsible for proper anteroposterior patterning, at least in bilateral animals (Finnerty & Martindale 1999; Finnerty et al. 2003, 2004; Wellik & Capecchi 2003). Additionally, cell-specific staining for the *Emx-Am* and *Pax C* genes showed a neuron-like ectodermal population that clearly resembled the cells identified when staining for GFP-like genes. Both genes play important roles in early neural tissue specification, as well as in eye morphogenesis (Miller et al. 2000; Ball et al. 2002; de Jong et al. 2006).

Considering the correlative spatial expression of key regulatory genes (see above) in a variety of cnidarians and the expression of GFP-like genes in *Acropora*, it may be that despite the apparently ambiguous physiological properties of these widespread genes/proteins (at least in the Scleractinia), their early and localised expression may indicate their involvement in tissue/cell characterisation, a potential role in chemosensory and photoreception (in the case of aboral expression; Pang et al. 2004; Chia & Koss 2005), and perhaps also symbiosis-related functions.

### 5.4.7 Cell surface array

The endodermal expression of both GFP-like genes made it impossible to observe ectodermal cell surface expression (*e.g.*, Figure 5.2, panel E). Fortunately, using the natural autofluorescence induced by the fixation process, this endodermal signal disappeared when the organisms were observed under the GFP2 filter, allowing the observation of the spatial pattern formed by ectodermal GFP-expressing cells (Figure 5.2, panel D). This aggregation characteristic has also been observed in the *Hydra* body column, using genes like *Hyzic* or *Nowa*, which are apparently expressed in proliferating stages by neuronal cell types, the nematocytes (Engel et al. 2002; Lindgens et al. 2004). In *Acropora*, however, GFP-like expression was not detected in that cell lineage, but was apparent in another neuronal cell type.
5.5 Conclusions

In *Acropora* embryogenesis, the GFP-like genes are maximally up-regulated during the planula stage, prior to metamorphosis. *In situ* and live fluorescent microscopy confirmed the expression of red and green fluorescent proteins at the aboral and oral ends of the planula, respectively. *AmilCP* was primarily distributed in the endodermal cavity.

In the ectoderm, two neuron-like cell types and one glandular-type cell showed associated mRNA and fluorescence signals. In the endoderm, GFP-expressing cells did not appear to be yolk cells, but rather may represent an as yet undescribed cell type.

*In vitro* assays of the temporal regulation of genes showed that *amilCP* was expressed before the *amilFP* gene during development.
6.1 Differential exon evolution, functionality, origins and regulators of the GFP-like genes.

During the process of sequencing the two GFP-like genes, the significance of protein structure for understanding the evolution of the corresponding genes became clear. The Acropora loci encoding both the GFP-like chromoproteins and their fluorescent counterparts shared a common intron/exon structure. When the Acropora’s GFP-like structure was compared with other nuclear homologues across cnidarians and bilaterians, the first two splicing cites do match extraordinarily well along the ORF in all organisms. In view of these similarities, it would be interesting to explore whether regulatory elements upstream and in introns are also conserved. The ability of putative regulatory elements to drive expression of reporter genes could be explored in well characterised mammalian systems, or in the emerging Hydra transgenic system (Bosch et al. 2002, Wittlieb et al. 2006, Khalturin et al. 2007).

To better understand the evolution of the coding region, amino acid sequences of GFP homologs were aligned and used to infer the phylogeny of the Acropora GFP-like proteins. The general topology of the resulting trees did not differ significantly from published trees, however, the most significant novel finding was that the larval proteins are largely resolved from their adult counterparts. This unexpected finding suggests that the adult and larval GFP-like genes may have different functions. It would be interesting to test this idea by gene silencing technology, which has been successfully applied in symbiotic anemones (Dunn et al. 2007), promising future applicability for coral embryos.

Finally, because the GFP-like proteins in general are capable of forming its chromophore, and some of them may be capable of light induced proton-pumping, it is possible to think of these proteins as enzyme like molecules. In terms of functional categorization, it was interesting to see that the complete protein sequence, and fragments (exons) hit the category of lipid and energy metabolism. However, the presence of several substitutions facing the outside of the B-barrel structure suggest
functions in protein-protein interaction, which could perhaps be explored using the yeast two-hybrid assay technology.

6.2 Protein world

It was fascinating to observe how bacterial colonies turn blue, or red after being transformed with expression constructs containing the coral GFP-like genes. However, the “cut and paste” did not always work exactly as expected. On occasion, the expressed proteins (particularly those in the RFP class) appeared to be toxic to the bacterial cell, as evidenced by poor growth and tendency to delete the plasmid under selective conditions. There are at least two possible explanations for this toxic behaviour. The first is that the self-aggregation and extraordinary resistant nature of the protein could lead to its physical disruption of sensitive structures inside the cell. The second potential mechanism of toxicity is more subtle, and is related to the secondary properties of the fluorescence, as has been documented in the case of killerRed (Bulina et al. 2006), which is derived from a hydrozoan cnidarian, highly reactive oxygen species may be generated as secondary products upon illumination. Further biophysical characterization of the Acropora RFP (amilRFP) promise interesting insights into biological functions of these proteins and their physical properties.

The AmilGFP517a (clone 114 c8m4) on the other hand, appeared to be only poorly soluble, but its spectral characteristics were almost indistinguishable (although slightly lower) from those of its paralog Amil517b (clone c012-c9). Similar low solubility phenomena were observed in the case of one chromoprotein. However by adding Triton X100 to 2%, the protein was fully solubilized. This behaviour was attributed to electrostatic changes on the surface of the protein. Whether such solubility behaviour applies in the coral cell remains unclear, however the majority of observed substitutions fall outside the barrel indicating that mutations in this latter protein domain may be strongly selected against.

The process of the acquisition of fluorescence by GFPS belonging to the non-fluorescent chromoprotein class has previously been reported (Gurskaya et al. 2001, Chudakov et al. 2003, Martynov et al. 2003), and occurs under strong light excitation. In the case of the Acropora chromoproteins, however, fluorescence is not acquired through light induction, rather dehydration causes the appearance of fluorescence.
This water-quenched fluorescence represents an important biotechnological characteristic of these proteins. Further studies on the crystal structure may reveal the mechanisms behind the conversion of non-fluorescent protein to the fluorescent state.

Finally, the spectral overlap between the GFP-like proteins, suggests that FRET may occur in living planulae, as has previously been demonstrated in adult coral tissue. (Salih et al. 2004, Cox & Salih 2005). The biological significance of these energy transfer processes is as yet unclear and, in the case of the planula may be restricted to the endoderm, since the GFP-like proteins have largely discrete expression domains in the ectoderm.

6.3 Too much attention to the core

Most structural studies on GFPs have focussed on the chromophore and the amino acids involved in its maturation (Chalfie et al. 1994, Cubitt et al. 1995, Gross et al. 2000, Gurskaya et al. 2001, Bulina et al. 2002, Campbell et al. 2002, Barondeau et al. 2003, Chudakov et al. 2003, Andresen et al. 2005). However, in terms of overall structure, the GFP proteins closely resemble the Nidogen (G2F) proteins. Whereas GFPs are classified as bioluminescent proteins in gene ontology, G2F proteins are categorized as constituents of the basal membranes (Ekblom et al. 1994, Kadoya et al. 1997). The significance of the overall structural similarity has not yet been explored, but lead to questions about the possible function of the ancestor of these very different proteins. The question is not easy to address; no GFP-like proteins have been identified yet in sponges or other “lower” animal phyla. The priority for future work should be to determine the expression pattern of the Acropora Nidogen-like gene (contig 1301). It is hypothesized that the Nidogen in situ pattern should overlap with those of the GFP-like transcripts, the logic being that it may be more difficult for an organism to create a new cell type which express the “derived” gene, than retain both messages in the same cell.

6.4 Environmental influences on early embryogenesis: does light controls GFP expression?

Light constitutes a form of energy that may be transduced by living organisms in a series of biological processes. The most obvious example is photosynthesis, the fixation of carbon dioxide driven by light energy. Another example is the circadian
clock, which is responsible for ensuring the correct metabolic state of organisms during active (growing, hunting) or inactive (sleeping) periods in terms of production/suppression of hormones, proteins and other secondary metabolites (McGinnis et al. 1994, Delaunay et al. 2000, Krishnan et al. 2001, Delaunay & Laudet 2002, DeBruyne et al. 2004, Kaneko & Cahill 2005). Light is also responsible of the evolution of complex mechanisms of photodetection and transduction in neural impulses, commonly called vision.

Each of these complex biological processes (photosynthesis, circadian clocks and vision) has two common characteristics, the first of which is that all are based on the absorption of light by a “chromophore” followed by processes of excitation and decoupling of the electrons excited or captured by this chromophore, generating an electrical potential that can then be used to produce “work”. The second characteristic common to each of these processes is the constancy of the input over evolutionary time, allowing organisms to adapt to the input.

Considering the evolutionary age and reactivity of the GFP chromophore, these proteins belong to a family with ancient origins, dating from before the split between the cnidarians and the bilaterians. A high proportion of the organisms which contain GFP-like proteins are regularly exposed to sunlight, as presumably were their evolutionary ancestors. On this basis, it is reasonable to postulate that some of the variation seen in coral GFP-like proteins reflects selective pressures for specific roles in light attenuation. However, this scenario may well not apply in the case of bioluminescent cnidarians, where different selective pressures may apply. If sunlight shapes in some degree the evolution of GFP-like proteins, then one might expect light to directly or indirectly regulate the expression of GFP-like genes. The AvicGFP for example, is capable of pumping protons upon light radiation, these protons could be sensed by other neighbour proteins and a network of activation or repression of GFP mRNA could be achieved. However, experimental support for this idea is lacking. In preliminary experiments, coral embryos raised in darkness showed relatively normal levels and distribution of mRNAs encoding GFP-like proteins (data not shown). Nevertheless, the notion of a link between light exposure and expression of GFP-like genes merits further exploration.
6.5 Future directions

Fluorescence Activated Cell Sorting could be used to isolate GFP/RFP expressing cells, allowing microarray profiling of the cell types. Electron microscopy suggest that treatment with a cocktail of digestive enzymes might be a more appropriate strategy to disaggregate larvae rather than mechanical disruption (Gates & Muscatine 1992). Cells released in this way could be cell sorted based on fluorescence characteristics, and then either be used directly in array experiments or used to found cell cultures.

Another important approach is that of examining the phenotypic consequences of knocking down the transcription of specific GFP-like genes during development. Although methods for knocking down coral genes are not yet available, morpholino oligonucleotide technology has been successfully applied to developmental genes in the case of the sea anemone *Nematostella vectensis* (Magie et al. 2007), and the hydrozoan *Clytia hemisphaerica* and *Eleutheria dichotoma* (Jakob & Schierwater 2007, Momose et al. 2008). Also RNA interference downregulation/knockdown has been applied successfully in the symbiotic anemone *Aiptasia pallida* (Dunn et al. 2007). These precedents offer great promise in terms of functional analyses of coral genes in general, and GFP genes specifically.