Final Consolidated Report.

Development of Mussel-Specific Expression Vectors Suitable for Transgenic Organisms: Implication on the Establishment of Mollusc Continuous Cell lines

Coordinator

[1] Claude Delsert
Laboratoire Génome et Populations
Station Méditerranéenne de l'Environnement Littoral.
1 quai de la Dorade. 34200 Sète
Tel: 33/4/67 46 33 90; Fax: 33/4/67 46 02 56; E-mail: cdelsert@univ-montp2.fr

Participant

Leonor Cancela da Fonseca
Universidade do Algarve
Unidade de Ciencas e Tecnologias dos Recursos Aquaticos
Campus de Gambelas
8000 Faro - Portugal
Tel: 351/89 800971; Fax: 351/89.818353; E-mail: lcancela@mozart.si.ualg.pt

[1]Present address: Centre de Recherche en Biologie Moléculaire - Centre National de la Recherche Scientifique - Route de Mende - 34293 Montpellier cedex5 Tel: 33/4/6761 33 30; Fax: 33/4/67 52 15 59; E-mail: cdelsert@xerxes.crbm.cnrs-mop.fr

Issuing date: July 1998.

CONTENTS

Summary

Executive Summary

SUMMARY

The main objective of this project was to develop the tools for the in-depth understanding of mollusc biology and pathology. Such a goal has been reached in other biological systems through the development of molecular biology and cell culture. Indeed, development of a mollusc cell culture system is an essential step to approach many of the problems related to mollusc pathology, in particular to use molecular biology tools to study viral infection of these organisms. No continous cell line is available from bivalves and while cell cultures have been obtained from these organisms they are difficult to develop and labor intensive, survival cultures being maintained only for a few weeks. Our goal has been to study in detail many of the parameters involved in maintaining primary cell cultures from bivalve molluscs in an effort to ease the burden associated with maintenance of these cells in culture. We then studied different ways to transform primary cell culture, in an effort to obtain continuous cell lines.

The initial part of this project was devoted to the development of mussel-specific expression vectors in order to help the transformation of mollusc primary cell culture. This approach required (i) the development of mollusc-specific expression vectors and (ii) the establishment of optimal conditions for primary cell cultures and (3i) transfection of expression vectors and detection of the expression of reporter genes and (4i) expression of oncogenes in primary cell cultures. A mussel genomic DNA library was built and was screened using the actin A3 gene of *Bombyx mori* as a probe. Clones corresponding to the mussel actin gene family were isolated and characterized. The promoter regions of five of these genes were subcloned and used to build mussel-specific expression vectors carrying the ß-galactosidase or Luciferase reporter sequence. In addition, the large T antigen of the simian SV40 virus, a potent oncogene in mammalian systems, was inserted in one of the expression vectors. Expression vectors were then transfected in primary cell cultures which were maintained in optimal culture conditions, as previously defined. Some success has been recorded as shown by detection of the reporter gene activity although the rate of transfection remained somehow low. Meanwhile, further experiments were performed (i) to determine transcription parameters such as the transcription start site and the tissue-specificity for each gene (ii) to analyze the level of expression of these genes at various developmental stages (iii) to look for cellular transformation by using carcinogenic drugs to treat different primary cell cultures or mollusc larvaes maintained in cell culture media (4i) to assess transient expression of a reporter gene by electroporation of axenic oyster larvae (5i) to uncover through a cDNA subtraction approach, a number of genes differentially expressed during *in vitro* cell culture as compared to gene expression in non-dispersed tissues.

Data concerning the molecular biology of molluscs and our observations regarding the obtention and transfection of mollusc primary cell cultures are discussed.

EXECUTIVE SUMMARY

The main objective of this project was to develop the tools for the in-depth understanding of mollusc biology and pathology. Such a goal has been reached in other biological systems through the development of molecular biology and cell culture. Indeed, development of a mollusc cell culture system is an essential step to approach many of the problems related to mollusc pathology, in particular to use molecular biology tools to study viral infection of these organisms. No continous cell line is available from bivalves and while cell cultures have been obtained from these organisms they are difficult to develop and labor intensive, survival cultures being maintained only for a few weeks.

The initial part of this project was devoted to the development of mussel-specific expression vectors in order to help the transformation of mollusc primary cell culture. The research programme contained the following main steps: (i) the development of mollusc-specific expression vectors, (ii) the establishment of optimal conditions for primary cell cultures, (3i) the transfection of expression vectors and subsequent detection of the reporter genes activity and (4i) the expression of oncogenes in primary cell cultures to help cell transformation.

Although a number of mollusc cDNA sequences have been reported, few reports concerned bivalves and yet none provided information about gene structure, especially about their promoter sequence. Therefore, the programme initial step was to build a genomic library with the DNA of blue mussel, *Mytilus galloprovincialis* and to isolate mussel genes and to analyze their structure.

Cloning of mussel actin genes.

A genomic DNA library was made by obtention of 9 to 23 kb-long DNA fragments made by the partial enzymatic digestion of high-length mussel-DNA fragments, then ligated in the phage Lambda Fix2. The resulting genomic DNA library contained 1.2 million independent clones. This number is equivalent to about 3 fold the number of clones required for a 99% chance of isolating an individual sequence from the full-length mussel genome (estimated to 1 billion base pairs), provided the DNA inserts have an average size of 15 kb. To identify and isolate mussel genes, a probe was initially made with a DNA fragment corresponding to the 5'end of the actin gene of the Lepidoptera *Bombyx mori*.

Screening was performed with the insect actin probe on 300,000 clones, a number of clones sufficient to provide a 99% chance of isolating any sequence of the genome and which is equivalent to 4.5 fold the length of DNA contained in the whole genome. Only 160 bp of the

400-bp-long actin probe overlapped the sequence coresponding to the N-terminal part of the actin coding sequence, the remainder corresponding to the 5'-untranslated region. Successsive rounds of screening were performed to isolate 32 distinct clones of which 19 were further characterized.

The average size of the library inserts was found to be about 14 kb, a size too large to be conveniently manipulated. Shorter DNA fragments corresponding to the targeted regions were then subcloned into plasmid vectors after identification according to the positive restriction subfragments identified after performing Southern blot analysis with the same insect probe. Sequencing of the positive subclones from three clones isolated from the library allowed the identification of an open reading frame encoding an amino acid sequence related to the well-conserved N-terminal part of the actin protein. These data allowed to identify two distinct actin genes, termed *mac*-1 and *mac*-2 and to design specific primers in the 5'-untranslated region. These gene-specific primers were used in a rapid PCR-test to sort the remaining clones. This analysis revealed that the pool of 19 initial clones contained 8 *mac*-1, 3 *mac*-2, and 8 actin genes of a different type.

Analysis of the 8 distinct clones through Southern hybridization with the same probe as above, subcloning and subsequent sequencing revealed 3 new actin genes: 2 *mac*-3, 1 *mac*-4 and 1 *mac*-5. The four remaining clones did not provide a clear hybridization signal and were not further characterized. Later on, while carrying out a population genetic study of the species complex *Mytilus,* we identified an additional actin gene by PCR amplification using degenerated primers on both genomic and complementary DNA (RT-PCR).

We thus have identified a family of 6 actin genes in *Mytilus,* a number which is comparable to what was observed in *Drosophila.* The different frequency of isolation of the distinct actin genes is likely to be inherent to the cloning procedure. It is likely that the incubation of genomic DNA with a unique restriction enzyme (*Sau*-3A) provided a slightly heterogenous pattern of digestion. The GC content is indeed known to be variable between distinct regions of the genome, thus leading to over- or under-digestion of these regions depending upon the sequence motif constituting the restriction site. For instance, we isolated either GC- or AT-rich microsatelitte sequences depending upon the restriction enzymes used when fully digesting the genomic DNA of another mollusc, *Ostrea edulis* (our unpublished observation).

Nevertheless, we believe that this genomic DNA library is representative of the mussel genome since we have been able to isolate clones each time we screened for various genes (our unpublished observation).

Sequence analysis of the isolated clones revealed they contained an open reading frame encoding for the first 42 amino acids and thus corresponding to the first exon of the actin gene. Such a structure is organized as in vertebrate actins, carrying a putative splicing donor site located at

the second nucleotide of the 42nd codon. In addition, a codon stop is found immediately after the last codon in all the mussel actin genes but *mac*-1 where it is located 23 codons downstream. The N-terminal sequence was found identical for *mac*-1 and *mac*-2 but distinct for the other genes. Alignment revealed that, as it is usually the case among invertebrates, mussel actin sequences are more related to the human cytoplasmic actin than to the skeletic actin. The N - terminus of all actins contains a series of 3 or 4 contiguous acidic residues between position 1 and 7 that constitutes the myosin binding site. Interestingly, *mac*-4 appears to be unique in that it contains an alkaline residue (lysine in position 4) in the series of acidic residues. Whether this has any significant biological effect is unknown. In addition, *mac*-5 sequence carries a serine (hydrophilic) in position 7 instead of the usual alanine (hydrophobic). *mac*-6 sequence comprises between 0 and 3 amino acid substitutions on the first 15 residues and it is noteworthy that the substitution in position 12 is unique since an aspartic acid (pKa: 3.8) is replaced with a histidine (pKa: 6).

Comparison of the 5'-untranslated region from genomic and cDNA sequences provided the evidence that the exon encoding the N-terminus of *mac*-6 is not the first exon but instead that it is preceded by a first short and distant non-coding exon. Interestingly, this gene was found to be present as at least 3 copies per haploid genome and as many as 5 distinct protein sequences were identified for this gene in *Mytilus* populations. More data will be required to determine whether the observed protein sequence polymorphism is a mere result of the presence of this gene as a multi-copy gene or if instead the polymorphism is observed because it provides adaptative advantages.

Transcription and 5'-end structure of actin genes

Determination of the transcription start site was performed using a combination of techniques such as primer extension, RNase protection assay and sequencing of 5' RACE PCR products. Results of these experiments show that the leader sequences are 56-, 66-, 60-, 42-, 3 4 - and 55-nt-long for *mac*-1 to -6, respectively. These data provide the approximate cap-site location for *mac*-1 to -4, which is about 25 nt downstream a putative TATAA box. By contrast no such informations were provided for *mac*-5 and *mac*-6 since the genomic region in front of their exon-1 are not cloned. The size of the intron-1 of *mac*-6 is not determined but partial sequencing revealed that the intron length exceeds 1.4 kb.

In an attempt to localize the exon-2 of mussel actin genes, the entire open reading frame of the human actin was used as a probe but it failed to hybridize to the genomic clones, thus indicating that the intron-1 for *mac*-1 and *mac*-2 is larger than 7 kb and 1 kb, respectively.

Sequence analysis of the region upstream to the exon-1 revealed a promoter-like region with putative TATAA box motifs and GC and CArG boxes which are well-known enhancer motifs in other eukaryotic organisms. GC box is the binding site for the sp1 transcription factor while

CArG box is the binding motif for the protein belonging to the Serum Response Transcription Factor (SRF) family. These enhancers are commonly found in the promoter region of actin genes and have been shown to participate to the early regulation of the alpha cytoplasmic actin gene in mammals. Interaction between SRF and other factors suggests a mechanism through which cells can coordinate intrinsec information on cell specificity with the response to external regulatory signals and a mechanism which can control cell-type-specific expression and differentiation in both a positive and negative manner. Together with the informations provided by the cap-site determination, these data demonstrated that the 5' region of *mac*-1 to -4 is the promoter region that had to be cloned in order to build mussel-specific expression vectors.

We further characterized the actin messengers and their level of expression in various tissues. Organs were dissected (including some of the surrounding tissues) and RNA was extracted as previously described by Chomczynski and Sacchi (1989) and then polyA selected.

Northern blot analysis revealed a 1.9-kb-long signal for all the actin genes. Quantitative PCR was performed on RNA extracted from tissues obtained after dissection and it revealed some degree of tissue-specificity in the expression of the distinct actin genes. Comparison of the level of expression of a single gene in the different tissues revealed that *mac*-1 and *mac*-2 patterns are similar although *mac*-1 is rather expressed in the mantle border while *mac*-2 is more expressed in the internal part of the mantle. *mac*-3, *mac*-4 and *mac*-6 are more expressed in the gill while *mac*-4 is the unique gene to be substantially expressed in the foot. *mac*-5 is more expressed in the adductor muscle and in the mantle border. More experiments are underway to better define the tissue-specificity of these genes. In addition, RNA was prepared from larvaes at different stages and determination of the stage-specificity of each actin gene is currently underway.

Intron-length polymorphism as a genetic marker.

While sequencing various clones of the *mac*-1 gene isolated from different individuals, we noticed nucleotide polymorphism in the intron-1 region. Further study brought evidence for an insertion/deletion event in the intron-1 region thus providing a valuable genetic marker usable through PCR analysis and potentially useful for the analysis of population structure of the genus *Mytilus*, *M. galloprovincialis* and *M. edulis*.

The length polymorphism reported here provides a new nuclear marker to discriminate between *Mytilus galloprovincialis* and *M. edulis* populations. This is an interesting feature in view of the thorny systematics of the genus *Mytilus*, of the process of introgression in areas where the two species hybridize and of the selective processes inferred to affect e.g. *edulis*-like genomes in some areas of overlap. We believe that our understanding of the geographic

7

structure and history of *Mytilus* natural populations will benefit from additional studies using the length-polymorphism at Locus *mac*-1 as a genetic marker.

The determination of the intron-1 sequence of *mac*-1 in *M. galloprovincialis* revealed the presence of a large (65 bp) insert in allelomorphs *mac*-1b and *mac*-1c. An octonucleotide that was present as a single element in the mac-1a sequences was as a direct repeat flanking the large insert in *mac*-1b sequences, and it was absent at the 5' end of the insert in *mac*-1c. It is thus tempting to propose that the octonucleotide was duplicated through the insertion process, as this has been reported for genetic mobile elements, that gave the allele ancestral to both mac-1b and mac-1c alleles. Subsequently, a short deletion immediately upstream the insert presumably deleted the 5' octonucleotide, which led to *mac*-1c. PCR-amplification using primers located in more conserved regions of *mac*-1 Intron-1 allowed the characterization of the three allelomorphs *mac*-1a, b and c. Some preliminary results on other non-coding regions of the *M. galloprovincialis* genome showed that length polymorphism is probably common in the non-coding DNA of this species, and can lead to the development of population genetic markers with similar properties (Our unpublished observation). The sequences of *mac*-1 haplotypes reported here, although limited in number, demonstrate at once that allelomorphs may consist of alleles whose divergence is probably very old (e.g. *mac*-1a), and also that identical alleles (*mac*-1c) occur in different species (*M. galloprovincialis* and *M. edulis*) probably as a result of recent introgression. Future in-depth phylogeographic studies involving the sequencing of large numbers of *mac*-1 alleles in samples of individuals from different *M. galloprovincialis* and *M. edulis* populations are expected to reveal the apportion of ancestral polymorphism vs recent introgression in the genus *Mytilus*. A more detailed report and discussion of this work has been published elsewhere (Ohresser *et al*, 1997).

Cloning of a series of mussel genes and cDNAs

Meanwhile, other DNA fragments were PCR-amplified from mussel cDNA using degenerated primers designed among phylogenetically conserved sequences. This alternative approach led to the identification of partial sequences corresponding to several mussel genes. Among the amplified sequences we identified cyclin B1, protein phi 1 and phi 2, rho oncogene, and more interestingly two types of proteins which are tissue- and/or development-specific transcription factors, such as a family of nuclear receptors and a POU protein.

The short DNA fragment encoding a phylogenetically conserved domain of the POU protein was used to screen the mussel genomic DNA library. It led to the isolation of several clones, one of which was subcloned and fully sequenced. It revealed an intronless structure as in higher vertebrates. By contrast, the coding sequence showed no homology with mammalian sequence outside of the conserved domain. In mammals, this POU protein was shown to be a transcription factor specific of neuronal tissues and its response element (DNA binding site) has been identified. This mammalian POU protein was found to activate a promoter carrying its response

element when transiently expressed. However, preliminary data show that expression of the mussel POU protein in the same context as above does not activate mammalian promoter. More experiments are needed to determine whether this absence of function is due to a different response element or to a diffferent cellular environment (Collaboration with Dr. M. Wegner, University of Hamburg).

More recently, RT-PCR was performed on mussel cDNA using degenerated primers designed on phylogenetically conserved domains of *snail* and *twist*, two homeo-domain proteins involved in the development of mesoderm. Sequence analysis revealed that their predicted amino acid sequences are homologous to vertebrate homeodomain. These DNA fragments were used to screen the mussel genomic library. Partial sequencing of the isolated clones confirmed they encode homeodomain proteins homologous to vertebrate *snail* and *twist*. More work is underway to characterize these mussel genes in order to determine the phylogeny of mesoderm (collaboration with Dr. M. Cassan, CNRS, Gif-sur-Yvette).

Construction of bivalve-specific expression vectors.

The final goal of this work was to built mussel-specific expression vectors in order to express oncogenes into transfected primary mollusc cell cultures. Gene-specific primers were used to PCR-amplify the 5'-untranslated region (over 1 kb) of the four actin genes (*mac*-1 to -4) for which the promoter region was characterized. PCR fragments were size-selected, purified and blunt-ended with Klenow polymerase. Fragments were ligated in the linearized pß.Gal.basic vector, a commercially available plasmid (Clontech, USA) that carries the ß *galactosidase* reporter gene and the SV40 termination signal. This cloning provided the *p.mac*-1 to -4 ß Gal vector series. To potentially increase the actin promoter activity, the SV40 enhancer element was inserted infront of the actin promoter of each constructs thus providing the *p mac*-1 to -4 e-ß Gal vector series. These two series of mussel-specific expression vectors carrying the ß-*galactosidase* reporter gene were built in order to test transfection efficiency and the sensitivity of different detection assays. In addition, the ß-*galactosidase* activity was to be tested after transfection with various mammalian (early promoters of cytomegalovirus and SV40, late promoter of human adenovirus) and insect (cytoplasmic actin A3 of *Bombyx mori*) expression vectors. Later on the SV40 early coding region was PCR-amplified and ligated in *p mac*-4 after deletion of the ß-*galactosidase* reporter gene. This was made in order to transiently express the SV 40 large T antigen, a potent oncogene commonly used to transform mammalian cells. Last, the *luciferase* gene, a most sensitive reporter gene, is currently being inserted into the expression vectors in order to improve the signal to noise ratio in detecting the reesults of transfection events.

Bivalve cell culture

Development of bivalve continuous cell line through the use of molecular biology tools is the ultimate goal of this programme. the initial step of this second part of the programme was to study in detail many of the parameters involved in maintaining primary cell cultures from bivalve molluscs in an effort to ease the burden associated with maintenance of these cells in culture. We then studied different ways to transform primary cell culture in an effort to obtain continuous cell lines.

Parameters for mollusc primary cell culture have been studied such as conditions of culture sterility, determination of the best type of cell tissue donor, composition of culture media and cell attachment conditions. Cells originating from mussel tissues and in particular from mantle (a tissue found to be a good cell donor) were consistently more difficult to disaggregate and to obtain as individualized cells; in addition, many cells are flagellated and therefore do not attach to the plate surface. Another aspect has been the difficulty to obtain immature organisms since in the Algarve coasts, the reproductive season is extensive and most of the mussels obtained have their mantle invaded by gonads, which induces a great heterogeneity in the cultures. We therefore concentrated most of our effort on oyster where we found that heart tissue is the best cell donor. In order to improve cell attachment, different types of cell culture dishes were assayed as well as different ways of coating the dishes prior to plating the cells but this did not lead to a significant improvement of cell attachment. Because of the lack of commercially available culture media, supplementation of 3x L15 media with various extracts or drugs was assayed and in some cases it increased cell survival but not cell spreading. Interestingly, it was found that mollusc hemolymph, algae extract and gonad extract improved the constitution of a cell layer. Indeed, results obtained with calf serum supplementation were consistently less good than with hemolymph, which remains the best supplement for these cells. These data confirm previous report by Odintsova and coauthors (1994) showing that calf serum cannot be higher than 2% without altering cell morphology and survival. We also tried other serum substitutes such as ultraser (Gibco) but this did not improve cell division or survival.

Transient expression in oyster cells

While studying transfection conditions, we used two different procedures in order to prepare cells of two distinct types: epithelial-like cells and cardiomyocyte-like cells which were prepared according to the procedure we have developed (this programme, reports 1 and 2). Both cell types were transiently transfected the different vectors prepared by participant 1 after introducing DNA into oyster cells and embryos either by lipofection or electroporation. We also assayed another invertebrate system as positive control (i) the A3lacz vector under the control of the actin A3 promoter from *Bombyx mori* in insect (Sf9) cells from *Spodoptera*

10

frugiperda and (ii) in vertebrate epithelial A6 cells from *Xenopus laevis*. interestingly, following lipofection, the insect expression vector A3lacz was found to be active in both vertebrate and invertebrate cells (insect and mollusc primary cells). In addition, it is noteworthy to observe that the A3lacz expression vector is functional when introduced into oyster embryos thus confirming the usefulness of this heterologous expression vector.

All vectors were transfected into Sf9 cells and as expected the best efficiency was found for the A3lacz vector for which the ßGalactosidase activity reached two orders of magnitude higher than the control. The next best performance was obtained with the CMV promoter (10% of the level of expression induced by the A3 promoter) while the mollusc-specific vectors did not show a significant level of expression in insect cells.

Because the major problem was the high signal to noise ratio obtained during detection of the reporter gene expression, we tried to increase the sensitivity of our detection assay by i) using a luminiscent enzymatic detection assay and ii) substituting the lacz with the most sensitive luciferase reporter gene. We also tested a number of constructs containing different promoter regions and/or derived from different mussel actin genes which were constructed by participant 1. Despite this effort, we were unable to improve the reproducibility of our detection assay; this is most likely due to a combination of a too low transfection efficiency and a degree of variability inherent to primary cell work as it has been described in previous reports.

Genes differentially expressed in cell culture conditions.

The lack of cell division in marine mollusc system may not be solely due to the lack of adequate cell culture medium adapted to the marine environment, but it may instead reflect the lack of basic knowledge of the cell biology of these organisms. In fact, little is known about mollusc cell biology including their interactions and their response to various stimuli and the factors controling their growth and development. To address this issue we tried to determine at the molecular level, some of the modifications sustained by the cells when placed in culture conditions. We used a cDNA subtraction approach (Vieto and Huber,1997; Kuang et *al.*, 1998). to clone some of the genes that are differentially expressed in *in vitro* cell culture as compared to whole tissue and to determine their RNA steady state level in response to different cell culture conditions. We identified a set of genes differentially expressed between the *in vivo* and *in vitro* conditions and we characterized the genes overexpressed in cell culture conditions. This study should provide molecular tools for a better understanding of the biology of these cells once placed in culture. In the course of this programme, while working on a primary culture of oyster larvae we isolated a eukaryotic cell growing in culture conditions. Sequence analysis of the 18s ribosomal RNA gene showed that it is phylogenetically related to a unicellular parasite of oysters. We are currently working on the identification of this potential

11

parasite and if indeed this is confirmed, we then already have molecular tools to study this parasite.

Thus, the initial part of the AIR programme was successful in providing the expression vectors necessary for the second part of the programme. On a longer term prospect, the programme allowed the uncovering of informations necessary to approach new aspects of the biology of molluscs. Among the outcomes are the cloning of genes involved in the development of the organism and/or of specific tissues such as neuronal or mesoderm tissue. These approaches are being pursued through collaborations with more specialized laboratories. More interestingly, sequences homologous to mammalian nuclear factors were identified that indicate that the neuroendocrine system is well developed in mussels. This particular research domain is vast and it will require a lot more research to figure out its degree of complexity when compared to vertebrates. Because the neuroendocrine system is involved in the control of development, growth and reproduction, the availability of molecular informations should induce a rapid development of this research area.

Chapter 1. Introduction

In the last two decades, cultures of marine bivalves have been repeatidly devastated by epidemic infections. Yet, little is known about the origin and even the pathogens involved in these infections (Laucker, 1983). In some instance, viral agents were identified through electron microscopy of contaminated organisms (Farley *et al*, 1972; Elston and Wilkinson, 1985; Nicolas *et al.*, 1992) but, due to the lack of a suitable homologous cell culture system, there has been little progress in the indepth knowledge of these pathogens. Indeed, little is known about the biology of mollusc pathogens including their mode of infection and their reproductive cycle. Obtention of mollusc continuous cell lines would greatly benefit to the overall mollusc biology as a similar evolution has been observed in other biological systems.

Due to our inability to propagate in laboratory conditions the pathogens affecting marine bivalves and in regard to their economical importance, it appeared relevant to pursue this effort through a new approach to the development of mollusc cell cultures. To this day, despite repeated efforts only one continuous cell line from a freshwater snail, *Biomphalaria glabrata*, was succesfully isolated (Hansen, 1974; Bayne, 1978). Strikingly, although media well adapted to the culture of organs and to primary cell cultures have been available for some time (Bayne, 1976), no attempt succeeded in getting spontaneous cell transformation from primary cell cultures despite the fact this approach had been successful in other biological systems and since proliferative disorders have been reported for various bivalves (Mix, 1976). Indeed, at least in one particular case, the hemocyte neoplasia of *Mytilus trossulus* (Farley, 1976), the disease has been well characterized and cultures of neoplastic hemocytes were assayed. This attempt did not succeed, most likely due to the fact that hemocytes result from a cell lineage which is not yet characterized.

Here we describe the use of molecular biology for a new approach to this old problem. This project was devoted to the development of mussel-specific expression vectors carrying oncogenes in order to help the transformation of mollusc cells in continuous cell lines (Jahn *et al.*, 1996; Merviel *et al.*, 1994). In parallel, a study was carried out to improve conditions for primary cell culture such as to better define culture media and to test various tissues, e.g. non-fully differenciated larval tissues. Subsequently, mollusc-specific expression vectors carrying reporter genes were used to define transfection conditions. In addition cell immortalization was assayed using chemical carcinogens on oyster trocophore larvaes.

In parallel, we tried to look at some of the molecular alterations sustained by mollusc cells in their effort to adjust to *in vitro* conditions. This was performed through a cDNA subtraction approach and led to the identification of a number of genes differentially expressed during *in vitro* cell culture as compared to whole tissue. We established a catalog of genes to study the

biology of mollusc cells in culture. This should provide useful information to apprehend the complexity of marine mollusc cells.

This report provides the details and analysis of the most significant results obtained during this programme.

Chapter 2. Materials and Methods

Cloning and characterization of the actin gene family

High-length genomic DNA fragments of the mussel *Mytilus galloprivincialis* were obtained essentially as described (Ausubel *et al.*, 1996). Briefly, a whole animal was frozen in liquid nitrogen and was ground in a tissuemizer and incubated in the digestion buffer (100 mM NaCl, 10 mM Tris.Cl, pH8, 25 mM EDTA, pH 8, 0.5% SDS, 0.1 mg/ml Proteinase K) at 50°C for 16 h. DNA solution was repeatedly phenol/CHCl$_3$ extracted and dialyzed for extensive periods of time in large volumes of TE buffer at 4°C. 400 µg of genomic DNA was partially digested with *Sau* 3A (0.02 units/ µg of DNA for 40 min), ethanol precipitated, resuspended in 400 µl of TE and loaded on a 10 to 40% sucrose gradient made in buffer A (1M NaCl, 20 mM Tris. Cl, pH 7.5, 5 mM EDTA) and centrifuged in a SW28 rotor at 20°C for 20 h. After analysis on agarose gel, fractions containing the 9 to 25 kb-long DNA fragments were precipitated and ligated in the Lambda Fix II phage vector according to manufacturer's recommendations (Stratagene Cloning Systems, USA). 1.2×10^6 independent recombinants were obtained that correspond to 4 times the number of inserts of 15 kb required for a 99% chance of isolating an individual sequence since the mussel genome is about 10^9 kb-long. 3×10^5 clones were screened with a 400-bp-long radiolabeled probe carrying 160 bp of the exon-1 of *Bombyx mori* actin A3 gene (Mounier et Prudhomme, 1986). Filter hybridization (5x SSC at 55°C) and washes (0.5 X SSC at 55°C) were conducted at low stringency. Inserts of a subset of clones were excised with *Not* and ligated in the *Not* I-linearized pBluescript plasmid (Stratagene). Inserts were analyzed by digestion with various restriction enzymes and were submitted to Southern blotting with the same probe as above. Only one restriction fragment hybridized to the actin probe of *Bombyx mori* as well as to the full-length human actin cDNA (data not shown). Positive fragments were subcloned in pBluescript vector (Stratagene) and were sequenced using a kit according to manufacturer's recommendations (Pharmacia Biotech, Sweden).

RNA extraction

Total RNA was extracted according to the single-step method by acid guanidine thiocyanate-phenol-chloroform extraction (Chomczynski and Sacchi, 1987). Poly A selection was performed with oligodT chromatography twice in batch and a third time in a column in order to minimize the elution volume. Poly A RNA was quantified by measuring the absorbance at 260 nm.

Riboprobes and RNase protection.

Riboprobes were transcribed from inserts PCR-amplified from the same individual that provided the RNA. The 5'-non coding region of each actin gene was PCR-amplified with specific sense primers with the following coordinates relative to the the initiation codon (the A of ATG being defined as +1) (-135; -116),(-172; - 153), (-117; -97), (-400; -381), (- 248; -228) for *mac-1* to *-5*, respectively and a 2 fold degenerated common antisense primer located at coordinates (+68; +46). PCR fragments were cloned in pGEM-T (Promega, Wi) and provided the pM1 to pM5 plasmids for *mac*-1 to -5, respectively. pM1 to 5 plasmids were linearized with specific restriction enzymes (*Nco* I for *mac-1*, *Spe* I for *mac*-2 and -3, *Rsa* I for *mac*-4, *Cla* I for *mac*-5) and allowed the *in vitro* transcription of 293-, 301-, 244-, 254- and 308 nt-long riboprobes homologous to the genomic sequence on 203, 240, 185, 178 and 256 nt for *mac-1* to *-5* transcripts, respectively. 0.5 µg of polyA selected RNA were hybridized to 10^5 cpm of gel-purified riboprobe. Riboprobes for *mac*-1, -2 and -3 were hybridized at 30°C for 16 h and digested with RNase One (Promega) for 1 h at 30°C. Because the riboprobes for *mac*-3 and -5 are A+T rich, hybridization was done in absence of formamide and digestion was performed with Rnase One or RNase T1 for 1 h at 20°C.

Primer extension

Gene specific antisense primers were designed at coordinates (-1; -27), (-1; -30), (-1; - 26), (-1; -30), (-1; -32) for *mac-1* to *-5* transcripts, respectively. DNA labeling was performed for 5 min with [α^{32}P] dCTP and later chased with 1mM of cold dCTP. 5 µg of total RNA was hybridized to specific primers for 90 min at 65°C in 15 µl of hybridization buffer (150 mM KCl, 10 mM Tris-Cl pH 8.3, 1 mM EDTA) and then ethanol precipitated. Pellets were resuspended in 20µl of reverse transcription mix [10 mM Tris-Cl pH 8.3, 50mM KCl, 5 mM MgCl2, 1mM dATP, 1mM dTTP, 1mM dGTP, 10 µCi [α^{32}P] dCTP (3000Ci/mmol), 0.15 mg/ml Actinomycin D, 13 units of AMV reverse transcriptase (Eurogentech, Belgium)]. Reaction was incubated at 42°C for 5 min and then it was adjusted to a final concentration of 1 mM dCTP and further incubated for 1 h at 42°C. For*mac*-4, the protocol was modified as follows. RNA was denatured for 5 min at 90°C in 7 µl of hybridization buffer in absence of EDTA and then transfered directly to 42°C and adjusted to 20 µl with reverse transcription mix and allowed to proceed as above.

Characterization of the actin cDNA 5' end

RNA was extracted from a single animal and cDNA was synthesized and ligated to adaptors using a kit according to manufacturer's recommandations (Clontech, CA). Specific antisense primers (5'-GTGAATTACCAATGCGGCTACATCA-3' for *mac-3*; 5'-CAACGCGGCTACGTCTTTATCGT -3' for *mac-5*) were designed based upon the nt sequence of a single allele of the XL1 individual. The hybridization temperatures, 66 °C and 68°C, were experimentally determined

16

for an optimal specificity of amplification of the 5' end of *mac-3* and *-5*, respectively. Tha antisense primers used for *mac-1*, *-2*, and *-4* were as described above, in the Primer extension section. After gel purification and a second round of amplification, a smearing band of about 50 bp was obtained and cloned in pGEM-T and sequenced. In order to minimize the risk of artefactual nucleotide substitution, the whole experiment was done in duplicate. Sequence carrying substitutions were defined as alleles only when found on several clones of the two distinct experiments.

Calibration of reverse transcription and PCR amplification

RT and PCR amplification were performed under subsaturating conditions. RT was performed using 1 μg of PolyA RNA with 40 units of AMV reverse transcriptase (Finnzyme) in a total volume of 20 μl of RT solution (100 mM Tris, pH 8.3, 80 mM Kcl, 10 mM MgCl$_2$, 10 mM Dithiothreitol, 4 mM sodium pyrophosphate, 1.25 mM deoxyribonucleotide triphosphates, and 50 μg/ml of oligo(dT)$_{15-18}$) and in presence of 2 μCi of [α^{32}P] dCTP. Reaction was done at 42°C for 1 h. Unincorporated radioactivity was removed by 3 ethanol precipitations in presence of 2.5 M ammonium acetate. Efficiency of the reaction, estimated by counting the amount of incorporated [α^{32}P] dCTP in an aliquote fraction and size-selection on an alkaline agarose gel, was similar for all the samples .

By subjecting parallel samples to increasing amplification (15, 20, 25, and 30 cycles) we determined that at 25 cycles, the amplification rate started to decline. We then performed a series of PCR amplifications by adding serial dilutions of cDNA synthesized at a linear point in the curve and using a limited cycle number (20 cycles). We observed a linear increase in PCR product amount over a 100-fold range of cDNA concentration (0.1-10% of the RT sample) (data not shown).

PCR assay

0.1-10% of the RT sample was added to a PCR (50 μl final volume) reaction (Eurogentec) with 5 μCi of [α^{32}P] dCTP (3000 Ci/mmol), 50 pmol of each primer, and 1 unit of Taq polymerase (Eurogentec). PCR cycles were performed as defined above with the following temperature profile : denaturation at 94°C for 1 min, primer annealing at 45°C for 1 min, and primer extension at 72°C for 10 seconds. Typically 1, 2 and 4 μl of the PCR sample were loaded on a 6% polyacrylamide gel, and the PCR fragments were visualized by autoradiography. Gene specific sense primers were the following with coordinates (relative to the initiation codon) between brackets: TTATTAACCCATTAACTACCGG (-57,-36), CAATACTCTAACAAGTTGAAAGG (-65, -43), GGAATAGAAAGGTTGTAACCAGC (-46, -24), TGTTGTGACAGTAATAGAAAGGTTG (-50, -29), GAGAACTGTAGGTGTAGACAATTG (-44, -21), GTCAGTATCGAACTTCAGCC (-28, -9) for *mac-1* to *-6*, respectively. The antisense primer was common to all genes and as defined above.

17

Primary cell culture and media.

After external cleaning the oysters, were opened and internal cavities was washed with filtered-sterile sea water; hearts from several individuals were collected and washed i n filtered sea water supplemented with 1% penicillin/streptomycin solution. Tissue was minced and further washed 4 to 6 times and homogeneized using a manual homogenizer fitted with a teflon piston and dispensed in 25 cm² flasks (Greiner). Cells were allowed to attach in a minimal amount of media and after 2 hours additional medium was added and then changed everyday for the first four days in order to prevent contamination. To this point cells were about 80 to 90% confluent, mainly with individual cells (approximately 10^6 cells per plate). Medium was renewed twice a week and cells survived for approximately 2 to 4 weeks without gross morphology changes. Afterward, morphological changes were observed (Figure 1-2) , with attached cells being increasingly sensitive to medium renewal while mortality greatly increased. No apparent cellular division was observed in the culture. At this time, a large percentage of the cells went to suspension. Trypan blue exclusion viability test and microscopic evaluation showed that these cells were alive.

While several culture media were satisfactorily assayed for the first 2 weeks of culture, the best results were obtained with filtered sea water supplemented with 20% of oyster hemolymph (a filtered mixture removed from pericard cavity, abductor muscle, mantle, foot and gonad) and containing 1% antibiotic mix.

Primary mussel cell cultures were prepared using 1 year old individuals.The animals were washed as above and the gonade infiltrated mantle was removed and washed in filtered sea water supplemented with 1% of penicillin/streptomycin solution. It was then minced and further washed for 4 to 6 times and was then resuspended in 1ml of filtered sea water with 0.012% collagenase and 0.012% of pronase, and was further incubated for 6 to 8 hours at 4°C. Digested tissues were then washed 3 times and dispensed in 25 cm² flasks (Greiner). After 2 hours of incubation at 25ºC, media was renewed. The best results were obtained with the same media as described for oyster cells. Media was renewed twice a week and cells were kept alive for several weeks. However individual cell attachement was not as good as observed for oyster cells. In order to split cells, individual cells were obtained through tissue disaggregation. In order to obtain an optimal tissue disaggregation, cells were treated with a trypsin solution (0.12%) diluted into a 35/1000 salt solution lacking calcium salts. Cells thus remained unattached during this treament. However, individual cells remain very sensitive to any enzymatic treatment, particularly mussel cells, and when applied this procedure led to a lower cellular viability. In addition, among the surviving cells a bigger proportion was unable to reattach to the dish, even in the presence of cell attachment factors such as poly-L-lysine.

Treatment with substances known to promote alteration in cell division

Treatment with Ethyl Methane Sulfonate (EMS): Primary cell cultures were treated for different time periods (ranging from 1 to 6 hours) with increasing concentrations of EMS, ranging from 0.5 to 32 ml of a concentrated solution of EMS (1.17 g/ml, Cat# M-0880, Sigma). Concentrations were chosen according to results previously described for yeast cell mutagenesis. Cells were rinsed twice with a 5%-sodium thiosulphate solution (to inhibit EMS) in sea-water and then incubated with regular cell culture medium at 25ºC. Treatment with Quercitin: Cells were treated for various time periods with quercitin which was previously dissolved in 0.025% DMSO and then diluted to its final concentration in cell medium.

Transfection

Cells from oyster heart were plated at 3 million cells per well, in 24 well-plates, and each experiment were performed in triplicate. Cells were transfected by lipofection using the methodology recommended by the supplier (DOTAP, Boehringer). The complex was prepared by mixing 5 µg of DNA diluted into 50 µl of Hepes (20mM) to 100 µl of Hepes containing 30 µg of Dotap. After an incubation period of 15 min, the mixture was added to 0.5 ml of culture media and dispersed onto the cells. The complex was removed after an additional incubation period of 24 hours at 25°C and fresh media was added to the culture. Following another 24 hours incubation period, transfected cells were assayed through *in situ* detection of ßgalactosidase activity and/or by determining the ßgalactosidase activity in the cellular extract.

Identification of the expression vectors: A3lacZ, the insect expression vector contains the lacZ gene under the control of the actin A3 gene promoter from *Bombyx mori* (Mounier et al., 1987).The p*mac*-1 to -4 mollusc-specific expression vectors contain either the lacZ reporter gene or the firefly luciferase gene under the control of different segments of the promoter regions of *Mytilus* actin genes. Mammalian expression vectors are under the control of viral promoters (RSV, CMV and SV40) and they contain the lacZ reporter gene.

Suppression Subtractive Hybridization analysis

For Suppression Subtractive Hybridization analysis, double stranded cDNAs were prepared by reverse transcription and then digested with *Rsa* I restriction enzyme. The cDNA chosen to be the tester cDNA was then ligated to specific adapters and hybridized in high stringency conditions with an excess of *Rsa* I digested driver cDNA. Specific tester cDNA fragments were then selected by PCR amplification with polymerase ADVANTAGE mix as described in the Subtraction cDNA kit (Clontech) using adapter-specific primers. Two rounds of hybridization and amplification were performed. Generated PCR products were randomly cloned in pGEM vector using the pGEM-T-Easy cloning system (Promega).

DNA Sequence analysis

Partial sequences of cloned PCR products were obtained by double stranded DNA sequencing using plasmid-specifics primers. Sequence data obtained was then submitted to databank search using the gapped Blast research tools. nBlast and xBlast programs were used for homology searches in nucleotide and protein data banks, respectively.

Northern and "Reverse-Northern" blot analysis

Cloned cDNA fragments obtained by Suppression Subtractive Hybridization analysis were gel purified and then used as probes to check the specificity of RNA accumulation of the respective genes. "Reverse-Northern" blot was obtained by performing a Southern blot with PCR amplified double stranded cDNA obtained under optimized conditions. Probes were labeled with $[\alpha^{32}P]dCTP$ (3000 Ci/mmol, Amersham) using the Prime-It II kit (Stratagene) and hybridized overnight in 6X SSPE, 0.5% SDS, 50% Formamide, 50 µg /ml Calf thymus DNA at 42ºC. Membranes were stringently washed in 0.1X SSPE at 65ºC during 30 min., and autoradiography performed at -70ºC with Kodak X-Omat AR film.

Detection of the ß-galactosidase activity

For *in situ* detection, culture media was removed, cells were washed once with sea-water, fixed in 0.5% glutaraldehyde (note that larvaes were fixed in 2% glutaraldehyde) solution for 15 min. and then incubated in PBS with 0.6mM chloroquine at 30°C for 45 min to inhibit endogenous lysosomal ßgalactosidase activity. Cell layer or larvae were then washed 3 times in PBS and stained with a solution containing 5mM K-ferricCN, 5mM K-ferroCN, 2mM MgCl2, 0.1% Triton X-100, 0.01% sodium deoxycholate and 1mg/ml X-gal at 37ºC for 24 hours. Cells expressing the ßgalactosidase activity appeared blue stained.

Cell extracts for quantitative assay were prepared according to the manufacturer's recommendations and were kept at -80ºC until processed. Detection was performed with the Luminescent ßgalactosidase Genetic Reporter System II (Clontech) with 20 µl of cell extract using a scintillation counter (Beckman) equiped with a single photon option. Total protein was determined using the Bio-Rad Protein Assay kit with 10 µl of cell extract. Results are expressed per mg of total protein.

Chapter 3. Results

Isolation and characterization of the *Mytilus* actin gene family

A series of 19 clones were obtained by screening a mussel genomic DNA library with a probe corresponding to the 5'-end of the actin gene of *Bombyx mori*. Sub-clones corresponding to 3 clones still hybridizing to the same probe were sequenced and analyzed. An actin encoding open reading frame was identified in all three clones (with two clones sharing the same amino acid sequence) which were referred to as *mac*-1 and *mac*-2 genes. *mac*-1 and *mac*-2 specific primers were used to screen the 16 remaining uncharacterized clones of which 8 were found to be distinct genes.

A weakly degenerated primer was then designed to hybridize into the actin ORF and was used to sequence these 8 distinct genes. Actin encoding sequences with amino acid substitutions, were found in 4 of these clones which were termed *mac*-3 (2clones), *mac*-4 and *mac*-5 while the 4 remaining clones were not further characterized. Therefore, the 19 initial clones led to the identification of 8 *mac*-1, 3 *mac*-2, 2 *mac*-3, 1 *mac*-4 and 1 *mac*-5 genes.

Full-length sequencing of *mac*-1 to -5 promoter regions was carried out and it determined the existence of a 126-nt-long exon with a distinct actin-related amino acid sequence for each gene. In addition the 5'-non-coding region was found different for each gene confirming the assumption that five distinct mussel actin genes were identified.

The 5'-non-coding sequences of the different actin genes were found to share no homology. Different transcription regulatory elements (enhancer), including GC-box, CArG-box, CCAAT-box, are similar to the elements known in vertebrate promoter regions. They are not present or located at the same coordinates in all the five genes (Fig. 1).

In addition, sequencing of two *mac*-1 promoter regions revealed the existence of *mac*-1 alleles which were used as genetic markers in the study of the *Mytilus* species complex.

Amplification of the actin cDNA 5'-end was performed by anchored-PCR using an antisense primer corresponding to a consensus sequence for the N-terminus of the mussel actin genes. DNA fragments migrating as 50 to 60 bp were obtained and their sequence corresponded to the previously identified genomic clones (data not shown). By contrast, a sequence distinct in both the leader and the amino acid sequence was identified and it provided evidence for an additional mussel actin gene, here referred to as *mac*-5 (Fig.1), for which we do not have a genomic clone. Interestingly, the cDNA 5'-end of *mac*-3 was not homologous with any of the actin sequences determined above, thus demonstrating there is a first non-coding exon before the sequence encoding the actin N-terminus. Because of its peculiar structure, this gene was set apart and thus is now refered to as *mac*-6. Southern blot analysis of genomic clones using the leader sequence of *mac*-6 cDNA as a probe did not reveal any positive clone (data not shown).
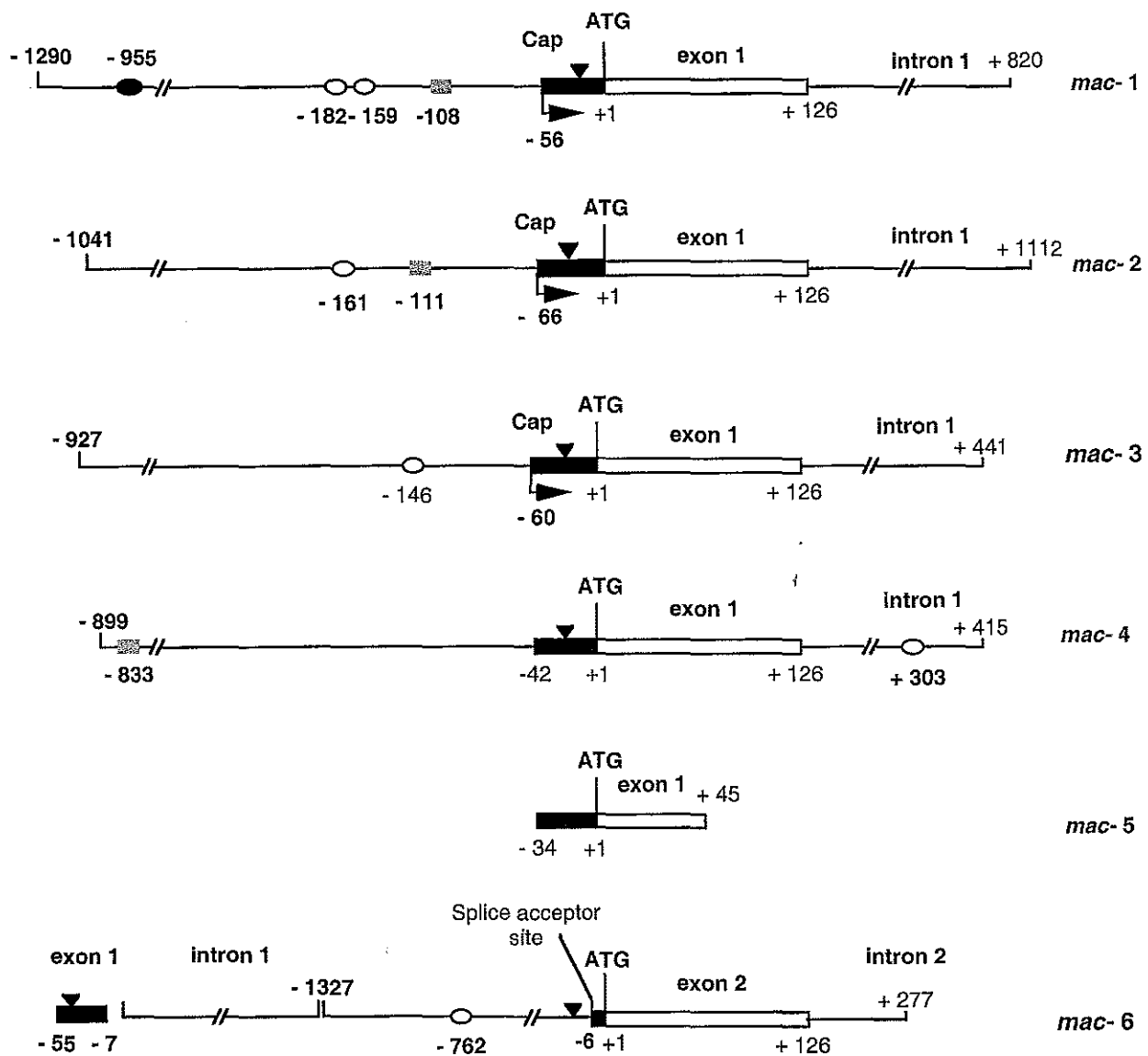
Fig. 1: Schematic representation of the 5'-end of the mussel actin gene family. Simple line for promoters and introns, open box for open reading frame, dark box for non-coding leader region, shaded box for GC box , dark oval for microsatellite sequence, open ovals for CArG box, dark triangle for homologous sequence of unknown function.Coordinates are in base pairs and are relative to the actin initiation codon with the A of the ATG being defined as +1.

Thus, the 55-bp-long cDNA leader sequence corresponds to a 49-bp-long first exon that is not present in the *mac*-6 genomic clones plus a 6-bp-long non-coding sequence of the second exon as indicated on Fig. 1.

Sequence analysis showed no homology between the different actin genes apart from a motif located in the leader sequences of each gene except *mac*-5 and it is also in the intron-1 of *mac*-6 (Fig. 1; Fig. 2). This motif is not significantly homologous to any other sequence and its putative function is unknown.

Analysis revealed that the predicted amino acid sequence for the exon-1 is identical for *mac*-1 and *mac*-2 while it differs by one or several substitutions with the other genes as summarized on Fig. 3. Conservation of the N-terminal part between the different *Mytilus* actins (88 to 100%) is not greater than it is with invertebrate or vertebrate actins (Table 1).

Downstream, the open reading frame is interrupted after the 42nd amino acid by a stop codon except in *mac*-1, itself preceded by a sequence homologous to the consensus for a splice donor site (Fig. 4). Actin introns at position 41/42 are conserved with chordates at the difference of lower invertebrates and arthropods which have few introns with variable positions (Sheterline *et al.*, 1995). No second exon could be localized in our genomic actin clones

Sequence polymorphism.

RNA and genomic DNA were separately extracted from single individuals. Anchored-PCR was performed using an antisense primer designed in the actin N-terminal amino acid sequence as above. In parallel, PCR amplification was performed on genomic DNA using the same antisense primer as above and sense primers specific for each actin gene (see Materials and Methods). DNA fragments amplified from genomic DNA extended from -116, -152, -380, -227, -96 to +45 (the A of ATG being defined as +1) for *mac*-1, -2, -3, -4, and -6, respectively. Sequencing of the DNA fragments amplified with *mac*-1 to -4 primers revealed two sequences corresponding to two alleles for each individual as it is expected for a gene present as one copy per haploid genome. For *mac*-6, five distinct sequences were found (data not shown) that correspond to at least 3 closely related genes per haploid genome. Allelic sequence polymorphism was found identical on both the cDNA and the genomic DNA, and amounted to 1.6%, 1.3%, 1.3% and 0.4% for *mac*-1, -2, -3, and -4, respectively, and it did not result in amino acid substitution. These values are in good agreement with data previously obtained for the intron sequence of *mac*-1 (Ohresser *et al*, 1997). By contrast, the percentage of substitution between the *mac*-6 sequences amounted to 5% for the intron-1 (4 sequences on 42 bp) and to 8% for the exon-1 (3 sequences on 64 bp). Interestingly, the five *mac*-6 sequences isolated from a single individual encoded 3 different sequences on the first 15 amino acids (Fig. 3, var2, var3 and var4) which are distinct from the original sequence isolated

Table 1. Amino-acids and nucleotidic pairwise identity score[a] in the 42 first amino acid region of the actin genes of *Mytilus galloprovincialis* and other species[b].

						Invertebrates						Vertebrate	
					Molluscs								
		Mussel											
	mac-1	mac-2	mac-6	mac-3	mac-4	mac-3[c]	A.c.n	A.c.m	P.m.	B.g.	A.s.	H.s.s	H.s.b
mac-1		100	93	93	93	95	100	98	98	98	100	86	95
mac-2	97		93	93	93	95	100	98	98	98	100	86	95
mac-6	89	89		95	95	88	93	90	90	95	93	88	93
mac-3	89	89	90		98	88	93	90	90	95	93	88	95
mac-4	89	89	88	97		88	93	90	90	93	93	81	93
mac-3[c]	ND [d]	ND	ND	ND	ND		95	93	93	93	95	86	93
A.c.n	79	82	78	74	73	ND		98	98	98	100	86	95
A.c.m	75	77	73	72	70	ND	87		100	95	98	83	95
P.m.	80	81	78	77	74	ND	85	82		95	98	83	93
B.g.	85	84	82.5	82	80	ND	82	80	82.5		98	88	98
A.s.	82	79.5	79	78	76	ND	84	81	82.5	78		86	95
H.s.s	68	67	68	63	64	ND	76	77	69	69	70		88
H.s.b	70	70	68.5	70	68.5	ND	82	84.5	83	77	77	79	
GC%	47	47	47	45	44	ND	60	63	59	53	53	69	68

[a] Noted in percent on the basis of Fig 3 aligment. Note that gap positions between pairs were not considered for identity score. Amino acid identity is above the diagonal, and nucleotide identity is under the diagonal. Coding sequence are determined from genomic lambda clones, excepted for mac-6 which sequence was deduced from a cDNA.

[b] For abreviation see fig 3 legend

[c] The undetermined 27 last amino acid of mac-6 region were considered as conserved

[d] Not determined

```
mac-1          -112    CTCCGCCC
mac-2          -115    CTCCGCCC

mac-1          -87     CATAAATA
mac-2          -99     CATATAAA

mac-1          -40                              AAAGGTTTTAAGACGTG
mac-1b1                                         AAAGGTTTTAAGACGTG
mac-1(XL1-1)                                    AAAGGTTTTAAGACGTG
mac-1(XL1-8)                                    AAAGGTTTTAAGACGTG

mac-2          -49                              GAAAGGTTCGC--ACGTG
mac-2(XL1-2)                                    GAAAGGTTCGC--ACGTG
mac-2(XL1-8)                                    GAAAGGTTCGC--ACGTG

mac-3            -57 AGTCAATTGTTGTGACAGTAATAGAAAGGTTGTTACAAGGT
mac-3(XL1-2)         AGTCAATTGTTGTGACAGTAATAGAAAGGTCGTTACAAGTT
mac-3(XL1-5)         AGTCAATTGTTGTGACAGTAATAGAAAGGTCGTTACAAGTT

mac-4          -29 AGACAATTGTTGAG
mac-4(XL1-22)      AGACAATTGTTGAG
mac-4(XL1-24)      AGACAATTGTTGAG

mac-6          -68             TGTAATAGGACCAGACAGGTTGCAAAAAATA
mac-6(XL1-2)                   TGTGATAGGACCAGAAAGGTTGCAATAAATA

mac-6(XL1-CL3) -55     GTTGTTTCAGGAATAGAAAGGTTGTAACCAGCT
mac-6(XL1-CL5)                 CAGTAA------GGTTGTTACAATCC
mac-6(XL1-CL6)                 CAGCAA------GGTTGTAACAATCC
mac-6(XL1-CL8)                 CAGCAA------GGTTGTAACAATCC
```

Fig. 2. Sequence conservation between different alleles of the different actin genes. Theses motifs are in the untranslated leader sequence of the actin RNAs except for *mac*-6 where it is also present in the intron-1 and for *mac*-5 in which it was not found. Nucleotide identity is shown in bold; coordinates are relative to the A of the actin initiation codon.

```
mac-1          MCDDD-VAALVVDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQ 42
mac-2          .....-....................................
mac-3          .D.E.-......I.............................
mac-4          .D.K.-......I.............................
mac-5          .....-LS........
mac-6 var1     ...E.-...M.I..............................
mac-6 var2     .....-.....IH...
mac-6 var3     .....-.....I....
mac-6 var4     .....-.........
mac-6 var5     ...E.-.....I....


P.m.           ....E-....................................
A.c.           .....-....................................
B.b.           ...E.-....................................
A.s.           .....-....................................
H.b            .-.E.-I...................................
H.a            ...E.ETT...C..............................
```

Fig. 3. Amino acid sequence comparison of the actin N-terminal part. A. c., *Aplysia californica* (neurons); P. m., *Placopecten magellanicus*; B. g., *Biomphalaria glabrata*; A. s., *Artemia salina* (actin 211). H. a, human α-actin (skeletic) gene; H. b, human β-actin (cytoplasmic) gene.

```
                                  126
                                   |
mac-1        CCC AGG CAT Cag gtA TGT GAA ATT 23 codons TAG
             Pro Arg His Gln Val Cys Glu Ile ...     ... Amb
mac-2        CCC AGA CAT Cag gtT CGT TAA TTT
             Pro Arg His Gln Val Arg Och Phe
mac-3        CCC AGA CAT Cag gtT TGT TAA TAA
             Pro Arg His Gln Val Cys Och Och
mac-4        CCC AGA CAT Cag gtT TGT TAA CAT
             Pro Arg His Gln Val Cys Och His
mac-6        CCC AGA CAT CaA gtT GGT TAA AAA
             Pro Arg His Gln Val Gly Och Lys
```

Fig. 4. Actin exon-1 splice donor site. Sequence alignment of the different actin genes showing a termination codon in position 45 except for *mac*-1 in which it is located 2 3 codons farther downstream. The splice donor site is in position 126. Small letters indicate the consensus sequence for splice donor site. Sequence for *mac*-5 is not determined. Coordinates are relative to the A of the actin initiation codon.

from the genomic library (Fig. 3, var1). An additional amino acid sequence (Fig. 3, var5) was identified on a separate individual.

Cap-site determination

In order to determine the cap-sites of *mac* genes and because of the allelic sequence polymorphism, probes were cloned from the genomic DNA of the individual that provided RNA. PCR amplification was performed on genomic DNA using the same antisense primer as above and sense primers specific for each actin gene (see Materials and Methods). DNA fragments extending from +68 to -135, +68 to -172, and +68 to -178 for *mac*-1, -2, and -3, respectively, were cloned and *in vitro* transcribed from pBluescript vector. RNase protection assay was performed using uniformly labelled riboprobes that provided 125-, 137- and 132-nt-long undigested fragments (data not shown) which correspond to 56-, 68-, and 64-nt-long leader sequences riboprobes for *mac*-1, -2, and -3, respectively (Fig. 1). Similar experiment was carried out for *mac*-4 in various conditions of stringency with RNase One or RNases A plus T1 and provided a protected fragment corresponding to the exon-1 length (data not shown). This result does not provide the *mac*-4 cap-site position but instead it indicates the location of an A/T rich region (a stretch of 17 A/T in position -15 to +2) which appears to interfer with the RNase assay. In addition this assay could not be done for *mac*-6 and *mac*-5 for which we do not have a corresponding genomic sequence.

Gene-specific primer extension using antisense primers starting 1 nucleotide before the initiation codon was performed in presence of $[\alpha^{32}P]$ dCTP. Reaction products were resolved on denaturing polyacrylamide gel and their size was precisely determined by the co-migration of a sequencing reaction. Elongation products of *mac*-1, -2, and -3 migrated as 56-, 66-, and 60-nt-long DNA fragments (data not shown) thus providing a cap-site location in agreement with the value obtained above by RNase protection assay.

Cap-sites are located 20 to 25 bp downstream a putative TATA box (7 A/T). Approximate location of the cap-site position for *mac*-4 to -6 was estimated by sequencing the longest products of anchored-PCR reactions performed in parallel. These experiments provided 52-, 34-, 55-nt-long leader sequence for *mac*-4, -5, and -6, respectively, as schematically represented on Fig. 1.

Tissue-specific expression of the actin gene family.

Northern blot analysis revealed a 1.9-kb-long signal for all the actin genes. Quantitative PCR was performed on RNA extracted from tissues obtained after dissection and it revealed some degree of tissue-specificity in the expression of the distinct actin genes. Comparison of the level of expression of a single gene in the different tissues revealed that *mac*-1 and *mac*-2 patterns are similar although *mac*-1 is rather expressed in the mantle border while *mac*-2 is

more expressed in the internal part of the mantle. *mac*-3, *mac*-4 and *mac*-6 are more expressed in the gill while *mac*-4 is the unique gene to be substantially expressed in the foot. *mac*-5 is more expressed in the adductor muscle and in the mantle border. More experiments are under way to better define the tissue-specificity of these genes. In addition, RNA was prepared from larvaes at different stages and determination of the stage-specificity of each actin gene is currently underway.

Isolation of a neuronal- and development-specific transcription factors.

PCR amplification performed on mussel cDNA using primers designed in a phylogenetically conserved region allowed the isolation of short sequences belonging to two distinct POU proteins. The sequence corresponding to a POU protein was used as a probe to isolate a unique clone from the genomic library. Sequencing of this gene revealed an open reading frame encoding a 432 amino acid long protein. This protein is homologous to the POU proteins of other organisms in the Homeo and POU domains. It otherwise displays no homology with other known proteins. Preliminary data indicate that this transcription factor is not able to transactivate mammalian promoters that are activated by the mammalian POU protein.

Degenerated primers were designed in phylogenetically conserved regions of two genes, *snail* and *twist* , involved in the development of mesoderm. PCR amplification performed on genomic DNA provided a sequence for both genes that were used as probes to screen the mussel genomic library.

This work allowed the isolation of 5 clones for each gene which are currently in the process of sequencing. Characterization of the tissue-specificity of expression of these two genes will be done in parallel with the POU gene.

A genetic marker for population studies in the marine mussels *Mytilus galloprovincialis* Lmk. and *M. edulis* L.

A novel intron length-polymorphism at the actin gene locus *mac*-1 is here reported and used as a genetic marker for population studies in mussels of the genus *Mytilus*. Two closely related genes, *mac*-1a1 and *mac*-1b1 from a genomic library of *M. galloprovincialis* were partially cloned and sequenced. They mainly differed from each other by a 65-bp insertion within their first intron. Polymerase chain reaction (PCR) primers were designed outside the insertion. The PCR analysis of 166 individual mussels from *M. galloprovincialis* and *M. edulis* populations revealed three size-classes of alleles or allelomorphs, two of which were of the expected sizes for *mac*-1a1 and *mac*-1b1. One allelomorph was absent from *M. edulis* samples, while it was present at substantial frequencies in *M. galloprovincialis* populations. The frequencies of the two other allelomorphs significantly differed between *M. galloprovincialis* and *M. edulis* populations. The comparison of six *mac*-1 intron sequences over 277 bp showed

at once that allelomorphs encompassed alleles differing from one another by substantial numbers of mutations, and that identical alleles were present in both *M. galloprovincialis* and *M. edulis* individuals, a probable result of the recent introgression between the two species. A very detailed report and discussion of this work is provided in annexe 1.

Optimization of the primary cell culture system

Most of the work has been focused on oyster heart that proved to be the best cell donor for our purpose. Oyster heart proved to be the most convenient source of cells and routinely provided uncontaminated, well-attached individual cells that formed a 80 to 90% confluent monolayer. By contrast, cells from mussel tissues and in particular from mantle (a tissue found to be a good cell donor) were consistently more difficult to disaggregate and provided small aggregates instead of individual cells. In addition, a great proportion were flagellated cells that would not attach because of their motility. Culture conditions to obtain attached epithelial-like cells have not changed from those described last year. We tried other serum substitute such has Ultraser (Gibco) without any better results in term of cell division or survival. Results obtained with calf serum supplementation are consistently worse than those obtained with hemolymph, which remains the best supplement for these cells. These results confirm previous data (Odintsova *et al*, 1994) showing that calf serum can be added up to 2% without affecting cell morphology or survival, but increasing its concentration is detrimental to the cell survival.

Characterization of a eukaryotic cell type from a culture of oyster embryos

Because embryos contain undifferenciated cells, it is an attractive source of potentially dividing cells. Attempts were made to obtain dividing cells from *Crassostrea gigas* embryos. Early stage embryos were obtained through *in vitro* fertilization, and were collected and enzymatically dissociated. Later stages such as trocophore and veliger stages were mechanically and enzymatically treated. Dissociated tissues were then spread in tissue culture plates and maintained in culture as described above in order to detect dividing cells. While many attempts failed to produce dividing cell culture, in one instance we observed apparent cell division. These cells were isolated and grown in culture at 28ºC using a sea-water based medium supplemented with L15 and serum substitutes. Its morphology clearly corresponded to eukaryotic cells (Figure 5) but did not to usual mollusc cells. The growth curves determined from the estimation of the number of cells on a Malassay counting cell indicated a doubling time of approximately 7 days. This culture was kept in culture for 24 weeks and succesfully sustained cryopreservation in liquid nitrogen. When thawed and placed in culture conditions these cells resumed their growth pattern. PCR analysis using *C. gigas*-specific primer pairs designed to amplify microsatelitte sequences (our unpublished data) did not allow the detection of the DNA markers specific for *C. gigas*. The 18S ribosomal RNA gene of these cells was partially cloned and sequenced. This gene shows some highly conserved nucleotide sequences

Figure 5 : Morphology of the unidentified cells isolated from cultures of oyster embryos.
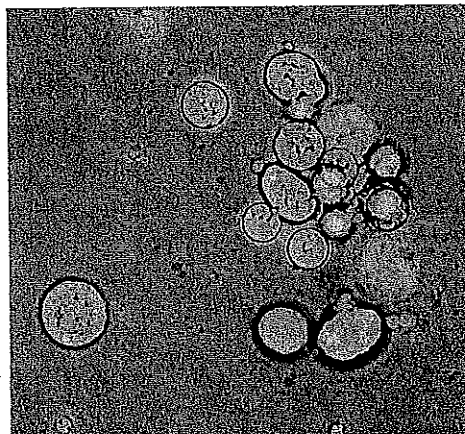
among all known eukaryotes and thus is conveniently used in phylogenetic studies. The corresponding region of *C. gigas* was studied in parallel as a positive control. Specific PCR primers were designed in a region of maximal conservation among eukaryotes. Forward primer starts at the 4th nucleotide of *Homo sapiens* 18S rRNA and the reverse primer was designed in order to PCR-amplify a 467 bp-long DNA fragment. Genomic DNA was prepared from *C. gigas* and from the unidentified cells. For both samples a 450 bp-long DNA fragment was obtained, as observed by agarose gel electrophoresis (data not shown). This DNA fragment was partially sequenced. As expected, the sequence derived from *C. gigas* shows a higher match with the 18S rRNA of two oysters, *Crassostrea virginica* and *Ostrea edulis* respectively. Instead, the sequence from the unidentified cells showed a higher match with *Developayella elegans* (Leipe et al, 1996), and with *Chaunacanthid* sp. (Amaral Zettler, Sogin & Caron, 1997) and with *Perkinsus* sp. (GenBank accession number L0737). *Chaunacanthid* sp.has been described as a "Single cell collected off the coast of Bermuda in September 1994", while *Perkinsus* is well-known unicellular parasite infesting some bivalves including *C. gigas.* In addition, histological procedures routinely used to identify *Perkinsus* failed to detect this parasite. Thus our results suggest that this organism is different from any known marine unicellular organisms. More work is being performed for a more precise identification of this organism.

Reporter gene expression in primary cell culture

We used two different preparation procedures in order to obtain cells expressing different phenotypes: epithelial-like and cardiomyocyte-like cells. Both cell types were transiently transfected with the different vectors prepared by participant 1 after improving the conditions to introduce DNA into oyster cells and embryos through lipofection and electroporation. We transfected by lipofection the following ßgal expression vectors: ßgal-mac-1 and pßgal-mac-2M and their derivatives pßgal-mac-2, pßgal-mac-1E and pßgal-mac-2E in the two types of oyster cell primocultures. The ßgal activity was detected *in situ* 2 days after transfection by Xgal chromogenic substrate (blue positive) in all samples (Table 2) except for non-transfected controls. Few cells (about10^{-5}) express enough ßgal activity to be detected by this qualitative method and inter-experiment variability was observed. Similar results were obtained when using 2 hour-old cultures after dissociation with collagenase in 25% hemolymph-containing sea water or when using 4 day-old cultures produced from mechanically dissociated cells maintained in 2X L15 medium with 2% fetal calf serum.

In order to increase the sensitivity of the essay and to obtain quantitative results, ßgal was enzymatically dosed (activity/total protein) in extracts of cells transfected with the expression vectors used above (Table 2) and with pßgal-mac-3, pßgal-mac-4 and pßgal-mac-5, (promoter coordinates are -964/-1, -893/-1, -899/-1 for *mac*-3 to -5, respectively, where +1 refers to adenine of translation start codon). More ßgalactosidase activity (mean values) was detected in cells transfected with any of the expression vectors than in non transfected cells (Fig 6). For both types of primocultures, we observed inter-

Table 2. Efficiency of the expression vectors derived from mussel actin genes (*mac*) in oyster heart primocultures.

Plasmid Name	Promoter		ORF	SV40 enhance	ßgal expression [b]
	gene	Coordinates [a]			
pßgal-mac-1	mac-1	-1069/-1	lac-z	-	+
pßgal-mac-1E	mac-1	-1069/-1	lac-z	+	+
pßgal-mac-2	mac-2	-746/-1	lac-z	-	+
pßgal-mac-2E	mac-2	-746/-1	lac-z	+	+
ßgal-mac-2M	mac-2	-990/-1	lac-z	-	+

[a] Relative to the A (+1) of the translation initiation codon. [b] *In situ* detection of the ßgalactosidase activity in transfected cells.
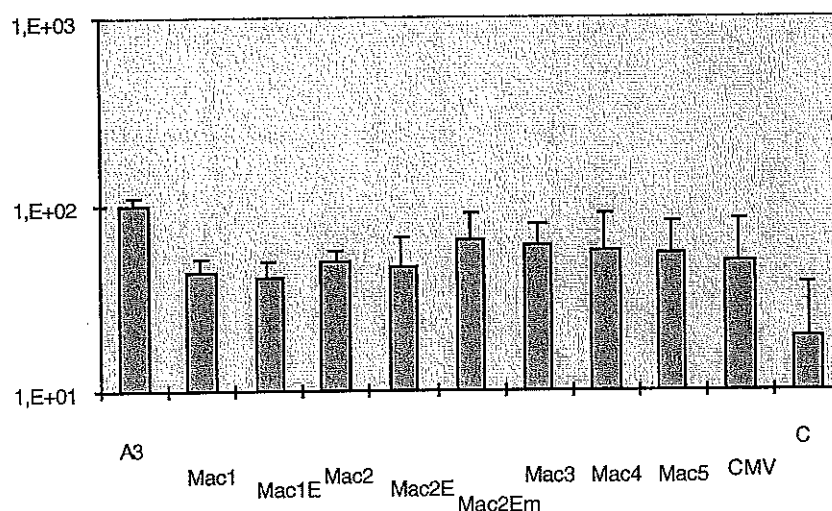
Figure 6: Transient transfection of the different vectors in oyster cells. Primary cultures of oyster cells from both epithelial and cardiomyocyte phenotype were transfected with the different constructs and the luminiscence catalyzed by the ßgalactosidase activity was detected as described above. Results of triplicate experiments are expressed as percent of the maximal levels of expression obtained for the A3lacZ vector (logarithmic scale), after being corrected for the total protein level of each sample. Identification of the plasmids is indicated under the bars and refers to those shown in Table 3. C for negative control.

experiment variability but also an unexpected high intra-experiment variability for which we do not have satisfactory explanation. Note that the heart cells form large clumps in culture conditions, which are sometime lost during transfection. Nevertheless, we conclude from these experiments that we were able to detect ßgalactosidase activity in the transfected cells since the analysis of variance shows that our results are significantly different from the control ($p < 0.01$).

The expression vector containing the promoter region of an actin gene from *Bombyx mori* (A3lacZ), which we have shown before to work well in cells from both vertebrate and invertebrate origin (see report 2), continued to be used as a positive transfection control both in a control cell culture system (Sf9, an insect cell line) and in the primary cultures of oyster cells. We also used this system to test the specificity of the mussel expression vectors. In addition we tried to improve the signal to noise ratio in our transfections results by increasing the sensitivity of the assay through the use of luminiscence activity, but we found no significant improvement suggesting that the limiting factor is not the sensitivity of the assay but the signal to noise ratio.

Treatment with drugs to promote alteration of cell division

A number of mutagens have previously been used successfully to induce increased cell division in a number of cell systems. The difference with our present cell culture system is that we have, at this time, no evidence that our cultures are dividing, thus making it more difficult to increase cell division by chemical treatment. We have tested two different mutagenic agents previously used with some success in other culture systems: (1) Ethyl Methane Sulfonate (EMS) which is a chemical routinely used to increase the frequency of mutants in yeast cells, randomly promoting GC to AT transversions (see *Current Protocols in Molecular Biology*, unit 13); (2) quercitin which is a flavonoid widely distributed in many edible fruits and vegetables, and it is one of the most potent mutagens found in bracken fern. Recent reports suggest that it may act in promoting mutagenesis and cell transformation, in particular when cells are both transfected with an oncogene and treated with this flavonoid. In our system, both substances either promoted high cell mortality or they failed to induce cell division in the cultures.

Transient transfection in axenic larvaes of oyster (*Crassostrea gigas*)

Artificial fecundation, is a commonly used technique that involves *in vitro* mixture of the two types of gametes. For this purpose, 3 female oyster were picked after gonad inspection. Gonads were striped and rinsed in 2 liters of filtered and UV-irradiated sea water. The same was done

Table 3. cDNAs differenctially expressed between oyster cell culture (C) and heart tissue (H)

cDNAs #	Related gene[d]	Identity score[c d]	Putative function	Reverse Northern	
				Heart/Cells[a]	Size
C1	eIF-4[A]	90% mammalian	Translation regulation	++/+++	1.8
C2 to C14	actin	99% oyster	Cytoskeleton compound	+/++++	1.3
C15	Plastin	64% mammalian	Cytoskeleton compound	+/++++	3
C16	Astacin	48% crustacean	endopeptidase	-/++++	2
C17	Rho-GDI 1	56% mammalian	RHO inactivation	+/++++	1.7
				+/+++	1.3
C18	Ribophorin I	64% mammalian	Protein glycosylation	+/++++	3.3
C19	AIF-1	59% mammalian	Allograff reject in mammalian	+/++++	1.6
C20	cytochrome P-450	48% mammalian	detoxification	+/++++	1.4
C21	EGF-like	34% *C. elegens*	ligand-receptor interaction ?	+/++++	1.7
H1	ADF	38% yeast	Cytoskeleton compound	++++/+	1.5
H2 to H3	MLC	78% scallop	Cytoskeleton compound	+/++	1.7
				+++++/++ +	1.2
H4	Collagen	42% mammalian	Extracellular matrix	++++++/+ +	1.4
H5	PTPase	34% mammalian	Phosphorilating membranar receptor	+/+++	1.6
				+/+	0.7
H6	CO I	60% echinoderm	Mitochondrial respiration	+/++	1.6
H7	NADH ubi	47% mussel	Mitochondrial respiration	ND[b]	
H8	16S rRNA	100% oyster	Mitochondrial tranlation	ND	

[a] Ratio of RNA accumulation between heart and cells cultured for 1 week in medium containing 80% hemolymph. [b] Not determined. [c] Sequences identification based upon the identity score. [d] See text for details

with a mature male gonad. Both gamete suspensions were filtered through 150 μm to eliminate debris. Suspensions were allowed to rest 30 min. for sperm activation and chorion brake down, and then mixed until each oocyte stayed surrounded by about 10 spermatozoids, and it was allowed to proceed for one hour. The eggs were then pelleted (2000 rpm/ 5min) and resuspended on 500 μl. 20 μg of DNA (A3lacZ) were added to 200 000 eggs and this suspension incubated for 5 min. in ice. Cells were then electroporated (1 pulse, 600V) in a BTX electroporator. Following electroporation, eggs were diluted 100 eggs/ml in sea water and egg development was allowed to proceed for 40 hours at RT°. Larvae were then fixed and expression of betagalactosidase was detected by histochemistry as follows: Larvaes were fixed in 1.25% Glutaraldehyde solution for 15 min, incubated in PBS with 0.6mM Chloroquine solution at 30°C for 45 min., and then washed three times in PBS buffer and stained with 5mM K-ferric Cn, 5mM K-ferro Cn, 2mM $MgCl_2$, 0.1% Triton X-100, 0.01% Sodium deoxycholate and 1mg/ml X-gal solution at 37°C for 24 hours. Data show that both trochophore and veliger larvaes were found expressing the characteristic blue staining indicative of ßgalactosidase expression while no staining was observed in control cultures. Our preliminary results showed a 1%-efficiency in term of transfected larvae. All larvaes were used at this stage so we have no information about the survival rate of electroporated larvaes.

Cloning of genes differentially expressed between whole heart and heart-derived primary cell culture.

Messenger RNA was isolated from both intact oyster heart tissue and from heart primary cell culture. Two subtractions of cDNA libraries were conducted using a commercially available kit (Clontech). These libraries were selected so that they would be enriched in genes preferentially expressed either in the donor tissues (heart) or in primary cell culture. Differential screeening of each library allowed the identification of clones overexpressed in each condition. Alignment of the partial sequences of the isolated cDNAs indicated significant homology with sequences in the data bank. 13 out of 32 initial clones were identified as actin genes, 8 more genes shared significant homology clones in the database (see Table 3) and the remainder showed no homology with any known gene. Since these clones were all obtained from subtraction library, they are all presumably expressed at high level under cell culture condition. Indeed, reverse Northern analysis confirmed that the isolated clones corresponded to genes with a higher RNA steady state level in cell culture condition and thus they correspond to genes moderately or highly expressed under cell culture conditions. This analysis indicated that the subtraction library was successful. The approximate cell to heart ratio of expression of these genes was estimated by comparing signal intensity between cells and whole tissue (Table 3).

Table 4. Comparison of the RNA steady state level for the genes preferentially expressed in heart primary cell culture when maintained in different culture media.

CDNAs #	Related gene	Size Kb	Sea water			2x Leibovitz medium	
			80% H[b] #1	25%H #2	2%S[c] #3	10%S #4	10%S+0.5%L[d] #5
C1	eIF-4A	1.8	+++	+	+	+	+
C3	actin	1.3	++++	++++	++	++	+++
C16	Astacin	2	+++++	+++	+	+	++
C17	Rho-GDI 1	1.7	++++	++	++	++	++
		1.3	+++	+	+	+	+
C20	P-450		++++	++++	+++	+++	+++
C21	EGF-like	1.7	++++	++	+	+	++
H1	ADF	1.5	+	+	+++	++	+
H4	Collagen	1.4	++	+	+	+++	++

[a]Estimated from Reverse Northern. [b]Hemolymph. [c]Fetal calf serum. [d]Lactalbumin hydrolysate.

Actin related cDNAs isolated from SSHc contained sequences with high identity scores (near 99%) with actin from *C. gigas* but also with vertebrate and invertebrate actin sequences. These partial cDNAs are either internal to the ORF (11 clones) or overlap the 5' untranslated region (UTR) (1 clone) or the 3' UTR (1 clone). The 5'UTR overlapping clone show sequence identity in the non-coding sequence with a previously reported *C. gigas* actin cDNA while the 3´ UTR overlapping clone (Cx) shows no match with this cDNA. So these 13 actin clones represent most likely transcripts from several different actin genes. When an actin coding sequence was used to probe reverse Northern blot, a 1.3 kb-long band was obtained. More actin RNA was detected in cells cultured with both hemolyph containing culture medium (Table 4) when compared to whole heart tissue (Tables 3 and 4). In addition, a small increase of RNA level was observed when lactalbumine hydrolysate was added to the medium (Table 4).

Chapter 4. Discussion

• The actin gene family.

Five actin genomic clones were isolated from a mussel genomic library made in the phage vector Lambda Fix. Sequence analysis revealed an exon carrying a 42 amino acid-long open reading frame homologous to the N-terminus of the actin proteins of other organisms.

The 5'-end gene structure. Analysis of the nucleotide sequence revealed an exon delimited by a sequence corresponding to the consensus for a splice donor site and a promoter-like region containing a putative TATAA box, and motifs homologous to the transcription enhancers of other eukaryotic organisms.

The most significant motif, the GC box that is the target of the sp1 transcrition factor, is present on *mac*-1, -2 and -4. Another element, the SRE element, is the target of factors belonging to the Serum Response Transcription Factors (SRF) family. These enhancers are commonly found in the promoter region of actin genes (Chen and Schwartz, 1996) and have been shown to participate to the early regulation of the alpha-actin gene in mammals. Interaction between SRF and other factors suggest a mechanism through which cells can coordinate intrinsec information on cell specificity with the response to external regulatory signals and a mechanism which can control cell-type-specific expression and differentiation in both a positive and negative manner (Gruenberg *et al.*, 1995).

Downstream, the open reading frame is interrupted after the 42nd amino acid by a stop codon except in *mac*-1, itself preceded by a sequence homologous to the consensus for a splice donor site (Fig. 4). Actin introns at position 41/42 are conserved with chordates at the difference of lower invertebrates and arthropods which have few introns with variable positions (Sheterline *et al.*, 1995). No second exon could be localized in our genomic actin clones.

PCR amplification of the actin gene 5'-end (RACE-PCR) allowed the identification of an additional actin gene refered to as *mac*-5 for which we have no genomic clone. The short N-terminal amino acid sequence and the short leader sequence clearly belong to a gene that is different of the previously identified mussel genes.

Two groups of actin genes. Analysis of the products of PCR amplification carried out on individuals revealed that five actin genes, *mac*-1 to -5, share the common following features of (i) being present as a single copy per haploid genome, (ii) displaying nucleotide sequence polymorphism at a rate of about 1% and (3i) encoding a unique protein sequence.

RACE-PCR provided evidence for an additional non-coding exon in one of the actin genes, termed *mac*-6. Its 49-nt-long first exon appears to be located far upstream since it was not found

in any of the genomic clones. This gene was found at least as 3 copies per haploid genome and it displayed quite a higher polymorphism than the other actin genes. In addition, at least five distinct protein sequences were identified for this gene in *M. galloprovincialis*. Such an amino acid sequence variability strikingly contrasts with the other mussel actins in which a unique sequence has been found. It can be hypothesized that this unusual sequence polymorphism results of the presence of multiple copies of this gene per genome, since it can be hypothesized that a single functional copy of a gene can complement for defective copies. This process would thus allow the accumulation of mutations in non- or less functional copies of the gene without the usual negative feedback by the pressure of selection. Properties of the mussel actin gene family is summarized on Fig. 7.

The N-terminus of the actin protein. Actin sequences are more highly conserved thatn almost any other proteins. Comparison of the actin sequence of a large number of organisms led to the finding that 13% of position in the amino acids are invariant and that 25 % of positions have substitutions of only one other amino acids (nearly always a conservative change). In addition, in 14% of positions this substitution occurs in only one actin sequence. For most positions amino acids are identical in the majority of actin sequences and in 248 positions 95% of the actins have the same amino acid. In this context, the N-terminal part is the most variable of the whole actin protein. This domain is acidic and it is involved in myosin activation and it is close or part of the binding site for a number of other actin-binding proteins (for review, Sheterline *et al.*, 1995)

The predicted protein sequences of the actin clones cover the first 42 residues except for *mac-5* for which only the first 15 amino acids were determined. The actin sequence is more homologous to the human ß-actin (cytoplasmic) than to the human skeletal α-actin as it is expected for invertebrate actins (Sheterline *et al.*, 1995). The actin N-terminal region contains acidics residues between position 1 and 7 which are involved in the binding to myosin and tropomyosin. This site, in most of the actins including 5 out of 6 mussel actins, is made of a series of 3 or 4 contiguous acidic residues (glutamic or aspartic acid). Interestingly, *mac-4* is the first actin reported to carry a lysine, a basic residue, in the myosin binding site (see Fig. 3). Whether this may induce a functional effect is not known, although it is important to underline the relatively high probability of apparition of a lysine in that context since a single nucleotide substitution is sufficient to change an glutamic acid in a lysine.

In addition, *mac-5* sequence carries a Serine (hydrophilic) in position 7 instead of the usual alanine (hydrophobic). *mac-6* sequence polymorphism goes from 0 to 3 amino acid substitutions on the first 15 residues and it is noteworthy that the substitution in position 12 is unique since an aspartic acid (pKa: 3.8) is replaced with a histidine (pKa: 6). More data will be required to determine whether the observed protein sequence polymorphism is a mere result of the presence

Fig. 7. Structure of the 5'-end of the Actin Gene Family of *Mytilus*

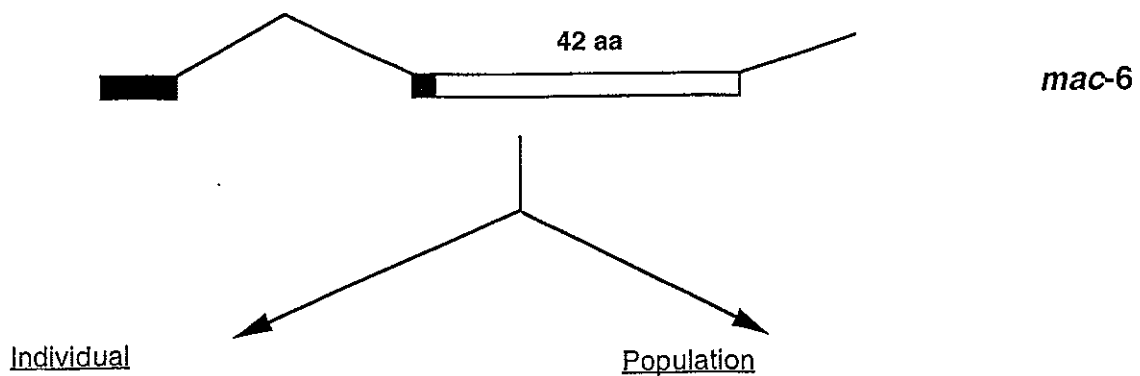## Group A: single copy genes

promoter   leader   exon-1   intron-1

42 aa
mac-1

42 aa
mac-2

42 aa
mac-3

42 aa
mac-4

15 aa
mac-5
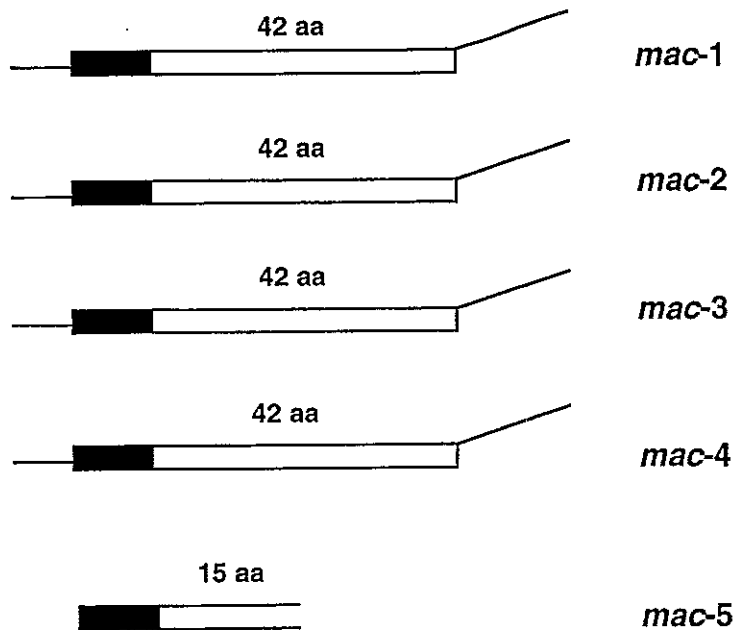
1 copy per haploid genome

A unique protein sequence per gene

Group B: multi-copy gene

exon-1   intron-1   exon-2   intron-2

42 aa

mac-6

Individual

3 copies per haploid genome
3 protein sequences per individual

Population

5 five protein sequences identified

of this gene as a multi-copy or if instead the polymorphism is observed because it provides adaptative advantages.

• Tissue culture from oyster larvaes

Our laboratory was not primarily involved in this part of the programme. Nevertheless, It was tempting to try to transform the undifferentiated tissues present in larvaes, despite the technical difficulties encountered. As mentionned above, it has been easier to obtain oyster than mussel larvaes and thus we attempted tissue culture and carcinogenesis on this organism. Two weeks old axenic larvaes were treated with EMS or with Dibenzo-carbazole and placed in culture media. They survived surprisingly well for days and then would sediment on the floor plate. Tissue would come out of the shell. Cells attached to the floor and after days would start to produce large spheric refringent cells. The most intriguing part of this series of experiments was the finding of a biofilm that developped in presence of various antibiotics and did not seem to interfer with the surviving tissue. Chemical transformation was also attempted on primary cell cultures (see results section) but did not succeed in inducing cellular division.

Lack of human ressources and cell culture facility, hampered this work, that could not be pursued. Nevertheless, we believe some aspects of this work should be more profoundly explored, especially the intriguing and possibly positive interaction between the contaminating prokaryotes and the cell culture.

• Isolation of developmentally regulated genes.

Although, blue mussel is a classical model in biology, study of more fundamental mechanisms suffered of the insufficient knowledge of its molecular biology. Soon after our first publication came out, we had contacts with research teams focusing on developmental biology for collaboration. This is the main reason why we isolated the homologue of a protein involved in the development of mammalian nervous system and why we more recently isolated two families of genes related to *twist* and *snail*, which are involved in mesoderm development in vertebrates. More demands are coming about other development-specific genes. This programme thus contributes to the progress of a dynamic scientific field. This interesting aspect promises to be the most immediate outcome of our programme.

• Characteristics of differentially expressed genes in oyster cell culture and whole heart tissue

Possible functions for the oyster genes identified by differential screening can be extrapolated from those of homologous genes extracted from the data bank and are summarized as follows:

DNA specific of heart derived primary cell culture. Actins, which are proteins of about 380 amino acids, are encoded by a family of genes (6 or more in mussel) differentially expressed in all tissues. These proteins are components of muscle fibers and microfilaments of cytoskeleton. Except for vertebrates and arthropods, identity of different actins genes cannot be discriminated by their amino-acid sequence.

Plastin/fimbrin genes code for 70 KDa actin-crosslink proteins of about 600 aminoacids (aa) implicated in the organization of the actin skeleton. Astacin family of metalloendopeptidase are extracellular proteins which all contain a N-terminal 200 aa-long protease domain with a zinc binding motif. Proposed functions of these proteases include activation of growth factor, degradation of polypeptides and processing of extracellular proteins. While crayfish astacin, a digestive enzyme secreted in digestive organ, and the fish choriolysins, implicated in hard chorion and egg envelope digestion, contain only the protease domain, other members of this family including Bone Morphogenetic Protein-1 (BMP-1) contain additional c-terminal domains. BMP-1 is a metalloprotease capable of inducing ectopic bone formation, which in humans, has a domain structure similar to that of the *Drosophila* dorsal-ventral patterning gene-product *tolloid*.

A common feature of proteins that include epidermal growth factor-like motifs is their involvement in extracellular functions such as adhesive and ligand-receptor interaction. Hybridization pattern obtained with *C.gigas* clones of astacin and EGF domain are very similar and are probably relevant of the same RNA containing both protease and EGF domains.

Translation initiation factor eIF-4A, a protein of 47 kD has been showned to carry a RNA helicase activity and it is implicated in regulating cell growth by selectively modulate the translation of different mRNAs based on their differing properties. Allograff inflammatory factor-1 (AIF 1), a small protein (17 kD), is expressed in macrophage of transplanted human heart and also in a rat model of chronic cardiac rejection. These proteins contain a calcium-binding domain and are related to calmodulin.

RHO GDP-dissociation inhibitor 1 (RHO GDI 1) is a 204 aa-long protein shown to inactivate RHO protein, which among other functions as been reported to be involved in cytoskeleton reorganization and cellular transformation. Ribophorin I is a 67 KD subunit of N-oligosaccharyltransferase catalyzing the transfer of a high mannose oligosaccharide onto asparagine within a consensus motif in newly synthesized proteins.

Cytochrome P450 belongs to a family of heme protein enzymes. These enzymes are widely distributed in animal tissues, plants and microorganisms and catalyze the monooxygenation of a vast variety of hydrophobic substances. They play an important role in the detoxification of drugs, mutagenes, and carcinogens.

cDNAs specific of heart tissue. Myosin interacting with actin filament is responsible of muscle contraction and cytoskeleton movements. Myosin is composed from 2 heavy chains and 4

light chains. Associated with each myosin head are 2 light chains, one essential light chain and one regulatory light chain with closely related sequences responsible for regulating the interaction of the myosin head with actin.

Actin depolymerizing factors (ADF) are proteins of 140-150 aa. These genes are differentially expressed and phosphorylated in different tissues, and play specific roles in regulating actin filament dynamics *in vivo*. Collagen proteins have been shown to form the major part of the extracellular matrix in many tissues. The transmembrane receptor protein tyrosine phosphatase (PTPase) sigma contain 2 intracellular PTPase domains and an extracellular region consisting of Ig-like and fibronectin type III-like domains, and is implicated in intracellular phosphorylation.

16S rRNA is a ribosomal RNA primarily transcribed from a large mitochondrial polyadenylated RNA. Cytochrome oxidase I (COI, 517 aa) and NADH-ubiquinone oxidoreductase chain 5 (441 aa in mussel) are mitochondria encoded proteins and components of the respiratory chain.

As expected, these results confirm a differential expression of genes between whole heart and heart-derived primary cultures, indicating that both cells types are adjusting to their environment. In both systems, genes involved in cytoskeleton reorganization were identified although not with the same specificity of accumulation. Interestingly, cells in culture overexpressed specific factors previously related to different stress conditions (RHO-GDI, AIF-I and cytochrome P450) as well as to development and transformation (RHO-GDI and Astacin-like protein), suggesting that cells are reacting to their new environment.

This response can be modulated by different cell culture condition as can be seen in Table 4, and is probably related to the different phenotypes observed after one week in culture.

As can also be seen in Table 4 the use of different concentration s of hemolymph, fetal calf serum and lactoalbumine hydrolysate, which are common additives to invertebrate cell culture media, can significantly affect the level of expression and/or the transcription of several genes in these cells.

In conclusion, our results show beyond any doubt that marine mollusc cells not only survive in culture but are able to response to external stimuli by adapting their transcription machinery to their needs.

References

Ali-Verma, T., Westphalen, P., Myllys, V., and M. Sandholm. 1997. Binding of *Staphylococcus aureus* to milk fat globules increases resistance to penicillin-G. J. Dairy res. 64(2): 235-260.

Amaral Zettler, L.A., Sogin M.L., and D.A. Caron. 1997. Phylogenetic relationships between the *Acantharea* and the *Polycystinea*: a molecular perspective on Haeckel's Radiolaria. Proc Natl. Acad. Sc. U.S.A., 94, 11411-11416.

Ausubel. 1996. Current Protocols in Molecular Biology. John Wiley and Sons, Inc.

Bayne, C. J. 1976. Culture of molluscan organs: a review, Invertebr.Tissue Culture.

Bayne, C. J., Owczarzak, and J. R. Allen. 1978. Molluscan (*Biomphalaria*) cell line : serology, karyotype, behavioral and enzyme electrophoretic characterization. J. Invertebr. Pathol. 32:35-39.

Chen C. Y., and R. Schwartz.1996. Recruitment of the *tinman* homolog *Nkx*-2.5 by serum response factor activates cardiac alpha-actin gene transcription. Mol. Cell. Biol., 16:6372-6384.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162:156-159.

Dalton, H. M., Poulsen, L. K., Halasz, P., Angles, M. L., Goodman, E. E., and K. C. Marshall. 1994. Substratum-induced morphological changes in a marine bacterium and their relevance to biofilm structure. J. Bacteriol., 176(22):6900-6906.

Delsert C., M. L. Cancela, M. Ohresser; and E. Lopes. 1996. Annual Report #1 "Development of Mussel Specific Expression Vectors suitable for transgenic organisms: implication in the establishment of mollusc continuous cell lines". European Contract AIR3.CT94.2487, EC (DG XIV).

Delsert C., M. L. Cancela, M. Ohresser, and E. Lopes. 1997. Annual Report #2 "Development of Mussel Specific Expression Vectors suitable for transgenic organisms: implication in the

establishment of mollusc continuous cell lines" European Contract AIR3.CT94.2487, EC (DG XIV).

Elston R, and M.T. Wilkinson. 1985. Pathology, management and diagnosis of oyster velar virus disease. Aquaculture 48: 189-210

Farley C.A., Banfield W.G., Kosnic G, and M.T. Foster. 1972. Oyster herpes-type virus. Science 178: 759-760

Farley, C. A. 1976. Ultrastructural observations on epizootic neoplasia and lytic infection in bivalve molluscs. Prog. exp. Tumor Res., 20: 283-294.

Flynn A, and C.G. Proud 1996. The role of eIF4 in cell proliferation. Cancer Surv. 27 : 293-310

Grueneberg, D. A., Simon, K. J., Brennan, K., and M. Gilman. 1995. Sequence-specific targeting of nuclear signal transduction pathway by homeodomain proteins. Mol. Cell. Biol., 15:3318-3326.

Hansen, E. L. 1974. A cell line from freshwater snail *Biomphalaria glabrata*. Int. Res. Comm. Syst., 2:1703.

Jahn, L., Sadoshima, J., Greene, A., Parker, C., Morgan, K. G. , and S. Izumo. 1996. Conditional differentition of heart- and smooth muscle-derived cells transformed by a temperature-sensitive mutant of SV40 T antigen. J. Cell Sci., 109: 397-407.

James, S.G., Holmstrom, C. and S. Kjellerberg. 1996 Purification and characterization of a novel antibacterial protein from the marine bacterium D2. Appl. Environ. Microbiol. 1996. 62:2783-2788.

Kelleher D.J., and R. Gilmore. 1997. DAD1, the defender against apoptotic cell death, is a subunit of mammalian oligosaccharyltransferase. Proc Natl. Acad. Sc. U.S.A. 13: 94 (10): 4994-4999

Laucker, G. 1983. Disease of mollusca : Bivalvia. *in* Kinne O ed. Disease of marine animal vol.II, Hamburg: Biologische Anstalt Helgland: 479-89

Leffers H, Nielsen MS *et al.*. 1993. Identification of two human Rho GDP dissociation inhibitor proteins whose overexpression leads to disruption of the actin cytoskeleton. Exp. Cell Res. 209(2): 165-174

Leipe D.D., Tong S.M., Goggin C.L., Slemenda S.B., Pieniazek N.J., and M.L. Sogin. 1996. 16s-like rDNA sequences from *Developayella elegans*, *Labyrinthuloides haliotidis* and *Proteromonas lacertae* confirm that the stramenopiles are a primary heterotrophic group. Eur. J. Protistol. 32: 449-458.

Lopes, E., Ohresser, M., Delsert, C., and M.L. Cancela. 1996. An insect expression vector capable of directing expression of the lacZ reporter gene in cells from marine invertebrate as well as vertebrate organisms. In Proceedings of the X Annual Meeting of the Portuguese Biochemistry Society, October 31 to November 2th, Braga, Portugal.

Lopes, E., Ohresser, M., Delsert, C., and M.L. Cancela.1997. Expression of Lac Z reporter gene under the control of an insect-derived actin promoter in larvae and primary cell cultures of oyster (*Crassostrea gigas.*) - In Proceedings of the III International Symposium on Research for Aquaculture : Fundamental and applied aspects. European Society for Comparative Physiology and Biochemistry. August 24-27, Barcelona, Spain.

Nicolas, J.L., Comps, M., and N. Cochennec. 1992. Herpes-like virus infecting Pacific oyster larvae Crassostrea gigas. Bull Eur Ass Fish Pathol 12(1): 11-13

Prassler, J., Stocker, S., Marriott, G. *et al*, (1997) "Interaction of a Dictyostelium member of the plastin/fimbrin family with actin filaments and actin-myosin complexes." Mol. Biol Cell, 8:1 83-95

Merviel, P., Degeorges, A., Salat-Baroux, J., and F. Calvo. 1994. Normal human endometrial cells in culture: characteriztion and immortalization of epithelial and stromal cells by SV40 large T antigen. Biol. Cell, 80:187-193.

Mix, M. C. 1976. A review of the cellular proliferative disorders of oysters (Ostrea lurida) from Yaquina Bay, Oregon. Prog. exp. Tumor Res., 20:275-282.

Odintsova, N. A., Ermak, A. V., and L. G. Tsal. 1994. Substrate selection for long-term cultivation of marine invertebrate cells. Comp. Biochem. Physiol., 107A (4):613-619.

Ohresser, M., Borsa, P., and C. Delsert. 1997. Intron-length polymorphism at the actin gene locus *mac*-1: a genetic marker for population studies in the marine mussels *Mytilus galloprovincialis* Lmk. and *M. edulis. Mol. Mar. Biol. Biotech.,* 6(2):123-130.

Sheterline, P., Clayton, J. , and J. C. Sparrow. 1995. Actin. Protein Profile. 2(1):1-68.

Stocker W, Grams F., Baumann U. *et al.* 1995.The metzincins-topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases Protein Sci. 4:5: 823-840

Utans, U., Quist,W.C., Mc Manus, B.M., *et al.* 1996. Allograft inflammatory factory-1. A cytokine-responsive macrophage molecule expressed in transplanted human hearts. Transplantation, 61(9):1387-92

Warbrick, E. and D. Glover. 1994. A drosophila gene encoding a DEAD box RNA helicase can supress loss of Wee1/mik1 function in Schizosaccharomyces pombe. Mol. Gen. Genet.245: 654-657.

Chapter 5. Finalization of tasks

• Programme Proposal

In the proposal for this research programme we proposed to develop the molecular biology of mussel in order to help with the transformation of mussel primary cell cultures in continuous cell lines.

The workplan included the following experiments:

1- Construction of a mussel genomic library and isolation of one or more genes.

2- Construction of one or more expression vectors.·

3- Establishment of primary cell cultures for *Mytilus*.

4- Expression of oncogenes in primary cell cultures.

5- Isolation of genetic markers.

6- Attempts to isolate mussel continuous cell lines.

• Achievements

1- Construction of a mussel genomic library and isolation of one or more genes.

The construction of the genomic library was successful since it produced a library with an unusually high titer and a good representation of the genome of *Mytilus galloprovincialis*. This is indeed proven by our success in isolating various genes from this library including the actin gene family, a neuron-specific gene (POU) and genes involved in the process of development (snail and twist).

Regarding the isolation of the actin genes we were primarily interested in the cloning of the upstream region of theses genes. We isolated a number of clones which allowed the identification of five distinct actin genes although they were not all characterized. This indeed might be the reason why we uncovered *mac5* only later on as a cDNA fragment and thus why we did not isolate a *mac*-5 genomic clone.

Because of the probe we used, we were successful in isolating clones containing the actin promoter regions except for *mac*-6 for which we isolated the region framing the first coding exon but not the promoter region because of an unexpected 5' non-coding exon farther upstream.

This first task was thus very successfully completed.

2- Construction of one or more expression vectors.

Expression vectors were build with the promoter region of the first four actin genes and also with the *mac*-6 gene before we discovered we did not have the promoter region for this gene. Nevertheless, the region of *mac*-6 we have cloned in a vector is still potentially a promoter since an actin gene has been reported in *Bombyx mori* with two promoter regions separated by a non-coding exon. Indeed, this region contains transcription regulatory elements. Two series of expression vectors were built with or witout the enhancer of SV40, a strong transcription activator in mammalian cells.

The ßgalactosidase reporter gene was inserted in these expression vectors. In addition, the most efficient luciferase reporter gene and the SV40 large T antigen, an oncogene, have been inserted in the *mac-4* expression vector.

3- Establishment of primary cell cultures for *Mytilus*.

Parameters for primary cell culture were studied for mussel and oyster. Oyster was found to be easier to manipulate and to routinely provide good quality primary cell cultures. Despite our attempts to ameliorate the culture medium with different nutrients and/or drugs and the culture conditions, we were not able to observe dividing cells. We thus believe that more research is needed to better define cell physiology and metabolism in order to understand at the basic level the absence of cell division in primary cultures.

5- Isolation of genetic markers.

While sequencing a number of clones of the *mac*-1 gene we uncovered sequence variability in the intron-1 region. This short region turned out to be a genetic marker that was used to analyse the *Mytilus* populations. *Mytilus* is a complex group of quasi-*species (M. edulis, M. galloprovincialis, M. trossulus)* that are closely related but still are distinct. These quasi-species sometime share the same geographic area and it is thus most usefull to have a molecular marker available. Indeed, this marker permits a clear distinction between *edulis* and *galloprovincialis* populations. A study based upon this marker has been published and more work is being carried out on larger geographic area.

4- and 6- Expression of oncogenes in primary cell cultures and to get continuous cell lines.

Experiments 4 and 6 are here discussed together since they are closely linked. At the difference of other biological systems (mammals, fish, batracians) no spontaneous cell transformation was ever reported in mollusc primary cell cultures. As initially mentionned in the programme proposal this part was the most hasardous of the whole project but it deservde to be tried for its major impact inncase of success. Here we can point out some of the reasons why we did not succeed in transforming cells.

First, in order to test our ability to transfect primary cell culture, we transfected by lipofection or electroporation, the different expression vectors carrying the ßgalactosidase gene as a reporter gene. Detection assay was performed *in situ* and it revealed the presence of a few blue-stained cells showing that the different expression vectors are functional but that transfection efficiency was limited. Conclusions were essentially the same when a luminescent assay was carried out instead of the *in situ* staining.

Nevertheless, because in theory a unique transformation event would be sufficient to provide a continuous cell line, an expression vector carrying the SV40 large T antigen (an oncogene) was lipofected in primary cell cultures; this assay did not succeed in inducing cell transformation. Therefore, beside increasing the transfection efficiency it might be necessary to try other oncogenes if possible cloned from molluscs.

As an alternative approach, oyster larvaes were used to as a source of dividing cells based upon the fact they contain non-fully differenciated cells. Young axenic oyster larvaes were placed, after incubation in a carcinogenic drug (EMS), in cell culture media where they survived for up to a week. They usually attach to the floor plate and a layer of cells most likely corresponding to the velum develops. This gives way to a brown tissue that survives for weeks. One problem we encountered was the presence of bacteria, some being inside of certain cells, that did not seem to be altering this cell culture. This observation raises the question of the relation of mollusc cells with the environment, including for some cells the possibility of symbiosis with bacteria.

We believe that more research is needed to better define cell physiology and metabolism in order to refocus this type of approach. In that prospect, an experiment was performed to identify genes that participated in the cell response to the culture conditions. Subtraction library was successful in cloning and identifying a number of genes overexpressed in cell culture. This work is in progress and it should contribute to our understanding of the cell adaptation and thus to the metabolism pathways affected in cell culture.

Experiments 4 and 6 are here discussed together since they are closely linked. At the difference of other biological systems (mammals, fish, batracians) no spontaneous cell transformation was ever reported in mollusc primary cell cultures. As initially mentionned in the programme proposal this part was the most hasardous of the whole project but it deservde to be tried for its major impact inncase of success. Here we can point out some of the reasons why we did not succeed in transforming cells.

First, in order to test our ability to transfect primary cell culture, we transfected by lipofection or electroporation, the different expression vectors carrying the ßgalactosidase gene as a reporter gene. Detection assay was performed *in situ* and it revealed the presence of a few blue-stained cells showing that the different expression vectors are functional but that transfection efficiency was limited. Conclusions were essentially the same when a luminescent assay was carried out instead of the *in situ* staining.

Nevertheless, because in theory a unique transformation event would be sufficient to provide a continuous cell line, an expression vector carrying the SV40 large T antigen (an oncogene) was lipofected in primary cell cultures; this assay did not succeed in inducing cell transformation. Therefore, beside increasing the transfection efficiency it might be necessary to try other oncogenes if possible cloned from molluscs.

As an alternative approach, oyster larvaes were used to as a source of dividing cells based upon the fact they contain non-fully differenciated cells. Young axenic oyster larvaes were placed, after incubation in a carcinogenic drug (EMS), in cell culture media where they survived for up to a week. They usually attach to the floor plate and a layer of cells most likely corresponding to the velum develops. This gives way to a brown tissue that survives for weeks. One problem we encountered was the presence of bacteria, some being inside of certain cells, that did not seem to be altering this cell culture. This observation raises the question of the relation of mollusc cells with the environment, including for some cells the possibility of symbiosis with bacteria.

We believe that more research is needed to better define cell physiology and metabolism in order to refocus this type of approach. In that prospect, an experiment was performed to identify genes that participated in the cell response to the culture conditions. Subtraction library was successful in cloning and identifying a number of genes overexpressed in cell culture. This work is in progress and it should contribute to our understanding of the cell adaptation and thus to the metabolism pathways affected in cell culture.

Coordinator   **Participant**

Claude Delsert   Leonor Calcela da Fonseca

Exploitation Report for the Programme AIR3 CT94-2487

Description of the results and innovative aspects.

The practical outcomes of this programme are:

1. A highly representative genomic DNA library of *M. galloprovincialis*.

This library was initially made of $1.2 \cdot 10^6$ independent recombinants. The master library which was built in the vector Lambda Fix 2 (Stratagene, San Diego, USA) was amplified only once and it contains $3 \cdot 10^9$ particule forming units per ml. The average insert-size was estimated to be 14 kb. This library is the first mussel genomic DNA library.

2. Sequence of the 5'-end of the members of the mussel actin gene family.

- The promoter, the exon-1 and part of the intron-1 regions of 4 actin genes (*mac*--1 to -4) were cloned and analyzed.

- The first non-coding and coding exons, part of the upstream and downstream region of the intron 1 and 2 of one actin gene (*mac*-6).

- The 5'-half of the exon-1 and the short leader sequence of one actin gene (*mac*-5).

All these sequences will be submitted to the genbank.

3. A series of mussel-specific expression vectors.

- p.*mac*-1 to -4 ß galactosidase vectors are under the controle of the actin 1 to 4 promoters and contain the ß-galactosidase reporter gene.

- p.*mac*- e. 1 to -4 ß galactosidase are similar to the above vectors but contain in addition the SV40 enhancer elements.

- p.*mac*- 4 T Ag vector contains the SV40 large T antigen, a viral oncogene widely used to transform mammalian cells.

- p.*mac*- 1 to 4 Luciferase vectors contains the Luciferase reporter gene under the controle of the actin 1 to 4 promoters.

These vectors are the first mollusc-specific expression vectors developped.

4. A genetic marker usable in population genetics.

An intron length-polymorphism at the actin gene locus *mac*-1 was identified and used as a genetic marker for population studies in mussels of the genus *Mytilus*. Two closely related genes, mac-1a1 and mac-1b1 from a genomic library of *M. galloprovincialis* were partially cloned and sequenced. They mainly differed from each other by a 65-bp insertion within their first intron. Polymerase chain reaction (PCR) primers were designed outside the insertion. The PCR analysis of 166 individual mussels from *M.*

galloprovincialis and *M. edulis* populations revealed three size-classes of alleles o r allelomorphs, two of which were of the expected sizes for *mac*-1a1 and *mac*-1b1. One allelomorph was absent from *M. edulis* samples, while it was present at substantial frequencies in *M. galloprovincialis* populations. The frequencies of the two other allelomorphs significantly differed between *M. galloprovincialis* and *M. edulis* populations.

This marker therefore allows a clear distinction between the two quasi-species *M. edulis* and *M. galloprovincialis.*

5. Development of an expression system for mollusc primary cell culture.

Transfection parameters were characterized for mollusc primary cell culture and embryo. Expression was obtained with both homologous (mussel-derived) and heterologous (insect-derived) expression vectors.

6. Isolation of a eukaryotic cell dividing in culture conditions.

While preparing primary cell culture from oyster larvaes, a dividing eukaryotic cell culture was isolated. Analysis of partial sequence of the 18S ribosomal RNA revealed these cells correspond to an oyster infecting parasite. This culture is preserved in liquid nitrogen and it will provide the molecular tools needed to study and diagnose this parasite in the wild.

7. A subtractive cDNA library from oyster.

A catalogue of genes over-expressed in whole heart tissues or in heart-derived cell culture was obtained that represent a series of molecular tools to a better understanding of the biology of these cells.

Description of the practical applications.

Results and tools provided by this programme have no immediate economical applications. Nevertheless, the practical outcomes mentionned above constitute a significant contribution to the knowledge of the biology of molluscs.

It is noteworthy to mention the scientific fields that already benefit or are likely to benefit of the results of this work:

- As mentionned above the isolated POU protein gene led to a collaboration with a lab specialized in molecular neurobiology (University of Hamburg).

- *twist* and *snail,* two genes involved in mesoderm development are being studied in collaboration with a lab specialized in the phylogeny of development (CNRS, Gif-sur-Yvette).

- The identification of a number of nuclear receptors and the many more clones waiting for sequencing show the importance of their participation into the physiological regulation in mussel. The impact of the understanding of their role in other biological systems show the potential benefit their study could bring to aquaculture because of their involvment in growth, development and reproduction.

- A list of oyster genes overexpressed in cell culture condition that will help to define the adaptation of cells . A collaboration on this project is under way with the Center for Marine Biotechnology at the UMBI in Baltimore (USA).

- Molecular tools to detect a single-cell organism dividing in culture conditions which is a putative oyster parasite. This study will be pursued as a collaboration with a marine pathology laboratory at the University of Porto.

- Mollusc-specific expression vectors are now available for more work on cell transfection of primary cell culture or mollusc larvaes. They have been carefully designed in order to facilitate their utilization. Convenient restriction sites were placed after the promoter and before the transcription termination site to facilitate the introduction of new protein coding sequences. In addition, we are currently defining their tendency in term of tissue-specificity, in order to optimize their expression.

Exploitation Plan

As mentionned above, this basic research programme was not intended to directly provide economical applications. It indeed intended to provide tools for the development of practical applications.
Accumulation of knowledge about the molecular biology of molluscs and the construction of series of mussel-specific expression vectors constitute important tools for further Research and Development efforts. Hopefully, these tools will help with the obtention of cell cultures provided a serious effort is maintained in that direction.

Publications

- Ohresser, M., Borsa, P., and C. Delsert. 1997. Intron-length polymorphism at the actin gene locus *mac*-1: a genetic marker for population studies in the marine mussels *Mytilus galloprovincialis* Lmk. and *M. edulis. Mol. Mar. Biol. Biotech.*, 6(2):123-130.

- Ohresser, M., and C. Delsert. Tissue- and development-specific expression of the actin gene family of the marine mussel. *Manuscript in preparation.*

- Lopes, E., Ohresser, M., Delsert, C., and M.L. Cancela. Transient expression of mollusc-specific expression vectors in oyster primary cell cultures. *Manuscript in preparation.*

- Lopes E., Ohresser M., and M.L. Cancela. 1998. "Identification of genes differentially expressed in primary cultures of heart-derived oyster cells (*Crassostrea gigas*)". *Manuscript in preparation.*

- Lopes E., Ohresser M., and M.L. Cancela. 1998. Identification of a a new parasite form from oyster embryos in Ria Formosa, portugal, and development of specific molecular tools for its study and diagnosis.)". *Manuscript in preparation.*

Supplementary investment

Dissemination activities

- Reports

- Delsert C., M. L. Cancela, M. Ohresser, E. Lopes. 1997- Annual Report #2 "Development of Mussel Specific Expression Vectors suitable for transgenic organisms: implication in the establishment of mollusc continuous cell lines" European Contract AIR3.CT94.2487, EC (DG XIV).

- Delsert C., M. L. Cancela, M. Ohresser; E. Lopes. 1996- Annual Report #1."Development of Mussel Specific Expression Vectors suitable for transgenic organisms: implication in the establishment of mollusk continuous cell lines" European Contract AIR3.CT94.2487, EC (DG XIV).-

- Delsert, C. 1996. 1er Rapport intermédiaire du programme AIR# 3 CT-2487 intitulé " Development of Mussel-Specific Expression Vectors Suitable for Transgenic Organisms: Implication on the Establishment of Mollusc Continuous Cell lines ". *in* Rapport d'Activité du Laboratoire GAP (IFREMER) pour l'Année 1995.

- Delsert, C. 1997. 2eme Rapport intermédiaire du programme AIR# 3 CT-2487 intitulé " Development of Mussel-Specific Expression Vectors Suitable for Transgenic Organisms: Implication on the Establishment of Mollusc Continuous Cell lines ". *in* Rapport d'Activité du Laboratoire GAP (IFREMER) pour l'Année 1996.

- Delsert, C. 1998. 3eme Rapport intermédiaire du programme AIR# 3 CT-2487 intitulé " Development of Mussel-Specific Expression Vectors Suitable for Transgenic Organisms: Implication on the Establishment of Mollusc Continuous Cell lines ". *in* Rapport d'Activité du Laboratoire GAP (IFREMER) pour l'Année 1997.

- Communication at conferences

Delsert, C., Lopes E., Cancela M.L.and Ohresser, M. 1998 - "Cloning of a mussel gene family and expression of vectors in primary cell cultures", 3th European Marine Science and Technology Conference - European Union participation, Lisboa. Portugal (Oral communication).

Lopes E., Ohresser M., Delsert C., and Cancela, M.L. 1998 - Identification of Genes Differentially Expressed Between Whole Heart Tissue and Heart-Derived Primary Cell Cultures from *Crassostrea gigas*. 3th European Marine Science and Technology Conference -May 23-27, Lisboa, Portugal (Poster presentation).

Lopes, E.,Ohresser, M., Delsert, C., Cancela, M.L. 1997 - Expression of Lac Z reporter gene under the control of an insect-derived actin promoter in larvae and primary cell cultures of oyster (*Crassostrea gigas.*) - III International Symposium on Research for Aquaculture : Fundamental and applied aspects, August 24-27, Barcelona, Spain - The European Society for Comparative Physiology and Biochemistry (Oral communication).

Lopes, E.,Ohresser, M., Delsert, C., Cancela, M.L. 1996. An insect expression vector capable of directing expression of the lacZ reporter gene in cells from marine invertebrate as well as vertebrate organisms. X Annual Meeting of the Portuguese Biochemistry Society, October 31 to November 2nd, Braga, Portugal (Poster presentation).