DTIC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

DTIC ELECTE AUG 0 1 1988 DULE

AD-A197 704

1b RESTRICTIVE MARKINGS: NA

3 DISTRIBUTION/AVAILABILITY OF REPORT: Distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): Boston University

5. MONITORING ORGANIZATION REPORT NUMBER(S): NA

6a NAME OF PERFORMING ORGANIZATION	6b OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Boston University	NA	Office of Naval Research

6c. ADDRESS (City, State, and ZIP Code)	7b ADDRESS (City, State, and ZIP Code)
Chemistry Department Boston University 590 Commonwealth Ave., Boston, MA 02215	800 N. Quincy St. Arlington, VA 22217-5000

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Office of Naval Research	ONR	N00014-86-K-0217

8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO.
800 N. Quincy St. Arlington, VA 22217-5000	61153N	RR04106	441c014	

11. TITLE (Include Security Classification)

(U) Characterization of Marine Bioadhesive Proteins

12. PERSONAL AUTHOR(S)

Laursen, Richard A.

13a TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Annual	FROM 7/1/87 TO 6/30/88	1988 June 10	3

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Adhesive, Proteins, Mussel, Amino Acid Sequence, Gene sequence, cDNA

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The objective of this research is to elucidate the amino acid sequences, via gene sequencing, of the adhesive proteins from several species of mussel and of other organisms, with the aim of understanding how these organisms attach themselves to wet surfaces. During the past year, we have cloned and sequenced fragments of the adhesive protein genes from three species of mussel. Two classes of protein are now apparent: they are similar in their content and location of lysine and tyrosine (or DOPA) residues, but different in repeat length and content of other amino acids.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	(U)

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
Dr. E. Smell, M. Marron or M. Haygood	(202) 696-4760	ONR

DD Form 1473, JUN 86 — Previous editions are obsolete — S/N 0102-LF-014-6603

SECURITY CLASSIFICATION OF THIS PAGE

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

R.A. Laursen
N00014-86-K-0217

1 July 1988

PROGRESS REPORT ON CONTRACT N00014-86-K-0217

PRINCIPAL INVESTIGATOR: Richard A. Laursen

CONTRACTOR: Boston University

CONTRACT TITLE: Characterization of Marine Bioadhesive Proteins

START DATE: 1 April 1986

RESEARCH OBJECTIVE: The primary initial objective has been to clone and sequence adhesive protein genes for several species of mussel with the aim of understanding what common (if any) structural features give these proteins their adhesive properties. It is hoped that this knowledge will lead to the development of adhesives that will have medical and other applications.

PROGRESS (YEAR 2): During the first year and continuing into the second, the focus of our work was isolating and sequencing several cDNA clones of fragments of the adhesive protein gene. This work showed that adhesive protein of *M. edulis* is primarily repeats of the decapeptide

xx1-Lys-xx2-xx3-Tyr-Pro-Pro-Thr-Tyr-Lys

where xx1 is usually Pro, Ser or Ala; xx2 is Pro, Ser, Leu, Ile or Lys; and xx3 is Thr or Ser. Using our original methods, however, we have not been able to obtain a clone or set of overlapping clones that encode for the entire protein. It appears that recombination, due to the repetitive nature of the gene, is occurring during cloning. Recently we have tried to overcome this problem we have fractionated our cDNA library, selected a fraction (3.3 kbp) large enough to code for the entire protein and are carrying out the subsequent cloning steps in recombinant-minus host strains. We have also isolated *M. edulis* genomic DNA and are currently screening the genomic library.

In year 2 we have also obtained sequence data from two other species of mussel, *Mytilus californianus* and *Geukensia demissus*, and we will soon have data from *Modiolus modiolus*. Cloning of *M. californianus* and of *M. modiolus* genes was carried out as for *M. edulis* by construction of a λgt10 cDNA library and screening with probes from *M. edulis*. The sequence of a clone from *M. californianus* was very simililar to that of *M. edulis*, except for the occurence of Arg (50% of the time) at position xx1 and about a 50% occurence of Ser and Ala at position xx7.

Cloning the G. demissa gene: A λgt10 library was initially constructed for this species, but screening with *Mytilus* probes was unsuccessful because (as we now know) of the significant sequence differences. For this reason, mRNA was isolated as usual and transcribed with reverse transcriptase to make cDNAs, which were then cloned into the LacZ gene the λgt11 expression vector. *G. demissa* adhesive protein was also isolated by extraction of phenol glands and purification by acid polyacrylamide gel electrophoresis. The protein containing band was excised and used directly to immunize a rabbit as a source of polyclonal antibodies.

The λgt11 library was packaged into phage and used to infect *E. coli* Y1090 host cells, and β-galactosidase fusion products were detected as colorless plaques among blue plaques in nonrecombinants. Colorless plaques were replated and screened with adhesive protein antibodies and an alkaline phosphatase linked secondary antibody, which gave a blue color with the substrate BCIP (5-bromo-4-chlor-3-indolylphosphate-p-toluidide). Positive clones were then sequenced in the usual manner. The *G. demissa* protein is significalntly different in that it contains repeats of from 11 to 13 amino acids, e.g.,

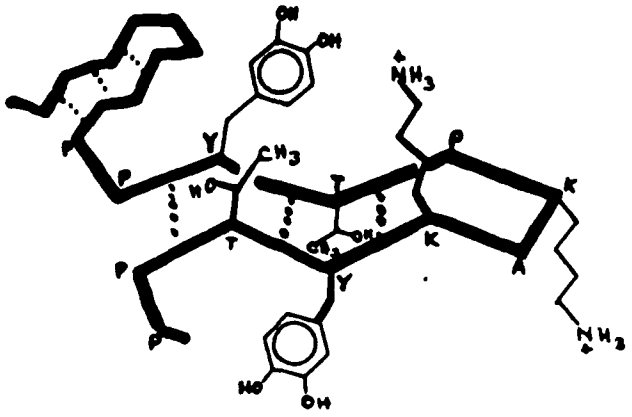
Gly-Lys-Pro-Thr-Thr-Tyr-Asp-Ala-Gly-Tyr-Lys-
Gly-Gln-Gln-Lys-Gln-Thr-Gly-Tyr-Asp-Thr-Gly-Tyr-Lys-,

and contains large amounts of glycine and glutamine, but little proline. Genetic material for this species was obtained by immunoscreening a λgt11 cDNA library, because, in contrast with the other species, we had no protein sequence data to guide the synthesis of oligonucleotide probes.

The recombination problem: Northern blot experiments in which a ^{32}P-labeled oligonucleotide probe is allowed to hybridize with mRNA have consistently shown, for all species, that the mRNA we have isolated from mussel phenol glands is long enough (3.0-4.0 kbp) to code for an adhesive protein with a molecular weight of up to 130,000. However screening of cDNA libraries shows not only many fewer clones than we might expect, but also much smaller, typically less than 500 base pairs, than the 3500 bp needed to code for the entire protein. This suggests that during the cloning process large amounts of information are being recombined out, even though we are using RecA$^-$ host strains. Furthermore, the fragments we have sequenced, and also those sequenced at Genex Corp. (unpublished), do not overlap, despite the fact that we have enough data to account for more than the entire protein. So the situation may be even worse than loss of information, there may also be scrambling.

Other workers have also encountered difficulty in cloning repetitious DNA sequences in certain *E. coli* strains. To overcome this problem, we are now beginning to carry out cloning operations in recominant-deficient hosts (RecA$^-$ and/or RecBC$^-$).

A conformational model for the Mytilus protein: Because of the invariability of Tyr and Lys residues and the patterns of posttranslational modification of Tyr and Pro residues, we believe that the adhesive proteins probably have some sort of regular, as opposed to a "random coil", structure. Given the large amount of proline, a structure with turns or loops seems more likely than a regular helical or sheet structure. Given the propensity for Tyr and Thr residues to occur in β-sheets, for Pro-Pro sequences not to be found in β-turns, but to cause a 90° bend in the peptide backbone, we have postulated the following β-sheet-β-turn model to serve as a working hypothesis for planned spectroscopic studies:



This model, though speculative, has some attractive features. It puts all the polar groups on the faces of the β-sheet loop, where they could interact with surfaces. In addition the Tyr and Lys residues are on both faces in pairs, in a symmetrical arrangement, where they might pair up with corresponding pairs in another chain to form interchain crosslinks. the major failing of this model is that one cannot make a similar model for the *Geukensia* protein, which contains little proline and has a less regular repeat structure. Of course *Geukensia* could have a completely different structure, but one would think, given the relatively constant placement of the critical Tyr and Lys residues, that there might be some conformational similarities. The answer to this dilemma can be answered only by experiment.

WORK PLAN (YEAR 3) In year three we plan to concentrate on obtaining the entire sequence of an adhesive protein either by sequencing genomic DNA or through the use of the rec$^-$ cloning strains mentioned above. Even if that fails, we now have or soon will have sufficient sequence data to begin analyzing the problem of what gives this class of proteins their adhesive character. During the next year we will focus on obtaining more sequence data on the *G. demissa* protein, because it is so different from the other species, and on getting sequence data from *M. modiolus*.

We plan also to begin conformational and modeling studies, using high resolution NMR techniques, on the proteins or peptide models, since it seems likely to us that these proteins have some sort of regular structure. If time permits, we hope to characterize the crosslink, which is presumed to occur between lysine and DOPA residues in these proteins, using chemical and mass spectrometric methods.