(14)

FINAL REPORT

SYNTHESIS AND CHARACTERIZATION OF HIGH MOLECULAR WEIGHT PEPTIDE POLYMERS AND COPOLYMERS CONTAINING L-DOPA RESIDUES

Contract No. N00014-86-C-0484

Biosciences Program
CORPORATE TECHNOLOGY

Submitted To:
Department of the Navy
Office of Naval Research
Attn: Procuring Contracting Officer, Code 151:MAW
800 No. Quincy Street
Arlington, VA 22217

July, 1988

AlliedSignal

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: U	1b. RESTRICTIVE MARKINGS: N/A
2a. SECURITY CLASSIFICATION AUTHORITY: N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): N/A	5. MONITORING ORGANIZATION REPORT NUMBER(S): N/A
6a. NAME OF PERFORMING ORGANIZATION: Allied-Signal, Inc. (6b. OFFICE SYMBOL (If applicable): N/A)	7a. NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code): Biosciences Laboratory, P. O. Box 1021R, Morristown, NJ 07960	7b. ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research (8b. OFFICE SYMBOL (If applicable): ONR)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-86-C-0484
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO: 61153N; PROJECT NO.: RR04106; TASK NO: NR441c019; WORK UNIT ACCESSION NO:

11. TITLE (Include Security Classification)
Synthesis of High Molecular Weight Peptide Polymers and Copolymers Containing L-Dopa Residues.

12. PERSONAL AUTHOR(S)
Williams, J.I., H.R. Bhattacharjee, I. Goldberg, A.J. Salerno, M.D. Swerdlof and P.D. Unger

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 8/86 TO 7/88		

16. SUPPLEMENTARY NOTATION
N/A

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP
08		

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
M. edulis polyphenolic proteins, L-Dopa, collagen, chemical peptide synthesis, enzymatic oxidation, recombinant DNA, adhesives.

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

The sea mussel M. edulis utilizes a polyphenolic protein in its adhesive plaque which contains tandem repeats of variants of a decapeptide sequence. We have chemically synthesized members of several families of peptides (GLUE peptides) related to a decapeptide sequence found in the M. edulis polyphenolic protein and prepared them in high purity for structure-function analyses. Chemical polymerization of model peptides and GLUE peptides using diphenylphosphorylazide (DPPA) activation has provided materials for lap shear adhesive strength tests. DPPA has also been used to prepare block copolymers between GLUE polypeptides and poly(ϵ-amino caproic acid). Concurrent enzymatic oxidation studies with GLUE peptides has given some insight into the crosslinking mechanisms which control relative reactivities of specific amino acid residues towards intramolecular or intermolecular bond formation. Polypeptides with repeating amino acid sequences have also been produced using recombinant DNA techniques. Several expression vectors based on the

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ OTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. M. Marron or M. Haygood	22b. TELEPHONE (Include Area Code): (202) 696-4760 — 22c. OFFICE SYMBOL: ONR

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

19. Abstract

regulatable p_L promoter have been developed and successfully employed in preparing synthetic gene cassettes for a collagen analogue, an elastin analogue, or a GLUE polypeptide. Gene stability and gene expression studies with a cloned collagen analogue gene have revealed several potential drawbacks to high level microbial expression of genes with repeating amino acid sequences.

FINAL REPORT

Date:	31 July 1988
Principal Investigator:	Jon I. Williams
Contractor:	Allied-Signal, Inc. P. O. Box 1021R Morristown, NJ 07960
Contract Title:	Synthesis of High Molecular Weight Peptide Polymers and Copolymers Containing L-Dopa Residues
Contract Number:	N00014-86-C-0484
Start Date:	1 August 1986

PROJECT SUMMARY

The research activities supported by this contract have been primarily devoted to preparing water-compatible adhesive peptides and polypeptides related to a sea mussel (Mytilus edulis) bioadhesive protein which contains L-Dopa residues. Subsidiary studies were undertaken within the scope of this contract to investigate mechanism(s) of oxidative curing of synthetic peptides containing tyrosine residues capable of conversion to L-Dopa or their derivatives with side chains. The results of these experiments and those determining lap shear adhesive strength of synthetic polypeptides containing oxidized tyrosines were used in concert with various spectroscopic analyses of synthetic peptides in solution to establish some preliminary empirical rules by which to judge the potential cohesive and adhesive character of polypeptide analogues to the M. edulis glue protein. Proprietary gene expression vectors were also developed to clone and express synthetic gene cassettes using recombinant DNA methodology. An important advance in the recombinant DNA research has been made in our studies of synthetic gene stability and stabilization of synthetic gene products to proteolysis. Synthetic gene cassettes for collagen, elastin, and M. edulis polyphenolic glue protein analogues were successfully cloned, and collagen-elastin analogue cocassettes were cloned in tandem as a model system demonstrating a biological route to the preparation of block copolymer polypeptides. Nylon-polypeptide block copolymers were prepared by chemical means in concurrent experiments and were

found to have thermogravimetric parameters that are generally typical of the Nylon polymer blocks.

INTRODUCTION

Protein glues have played a major role in adhesion technology for many years despite a lack of thorough understanding of the interfacial phenomena responsible for strong protein-substrate adherence. The costly problems of biofouling on exposed surfaces of ships, oil rigs, underwater pipelines and sea water intake valves for electrical power plants have stimulated a number of biochemical investigations into the nature of adhesion to solid substrates by marine organisms. The general hope has been to determine the molecular basis of recognition, attachment or adhesion to solid substrates for marine organisms as a prerequisite for designing antifouling agents. Such agents would represent a new generation of inhibitory or toxic materials that could replace the environmentally hazardous paints and primers presently used to discourage marine growth on submerged surfaces.

J. Herbert Waite and Marvin Tanzer [Science 212, 1038-1040 (1981)] purified a polyphenolic protein fraction from the Eastern sea mussel Mytilus edulis and demonstrated by amino acid analyses of acid hydrolysates the presence of L-Dopa (3,4-dihydroxyphenylalanine) and hydroxyproline residues in this material. These amino acids are seldom found in natural proteins, and are generally confined to certain structural proteins such as collagen and elastin and some of the exoskeleton proteins of certain invertebrates. Although the polyphenolic protein from the adhesive plaque which terminates the attachment byssus threads for M. edulis may not necessarily be the only adhesive substance in the adhesive plaque, it does tightly adsorb to a number of surfaces and is a prime candidate for being a water-compatible protein bioadhesive. Waite and his coworkers [Biochemistry 24, 5010-14 (1985)] have gone on to identify a number of biochemical characteristics of the M. edulis polyphenolic protein, including obtaining partial amino acid sequence information and examining its susceptibility to enzymatic oxidation by a mussel polyphenol oxidase. An interesting result from Waite's laboratory has been the observation of a large number (75-80) of tandem repeats of a decapeptide sequence within the mussel polyphenolic protein. Partial DNA sequencing of cDNA clones (prepared from sea mussel mRNA by the laboratory of

Richard Laursen at Boston University) has confirmed the predominance of decapeptide repeats within the gene or genes encoding the polyphenolic protein, and has shown that certain residues within these repeats (notably Lys-2 and Lys-10, and Tyr-5 and Tyr-9) are invariant from one repeat to the next.

It is not clear how the sequences and structural features of the sea mussel polyphenolic protein may directly contribute to the adhesive character imparted to it. Charge effects and hydrogen bonding capacity brought about by the density of hydroxylated polar side chain groups in the polyphenolic protein certainly can contribute to the free energy of adhesion for the mucoid adhesive plaque in which it is embedded. Waite [Int. J. Adhesion and Adhesives 7, 9-14 (1987); Chemtech 692-697 (November, 1987)] has drawn special attention to the presence of L-Dopa residues in the polyphenolic protein because of their high capacity for metal ligand chelation, their ability to displace water molecules from hydrophilic surfaces, and their capacity for oxidation to o-quinones which can further react with nucleophiles to form covalent bonds. The formation of covalent bonds can lead to increased adhesiveness or increased cohesiveness for the biomaterials of the adhesive plaque depending upon the nature of the bonding moiety. The pyrollidone residues proline, 3-hydroxyproline and 4-hydroxyproline found in the polyphenolic protein also offer further opportunities for hydrogen bonding while potentially promoting maximum adsorption to surfaces by introducing kinks in the protein backbone which favor an extended or open conformation. The presence of numerous decapeptide repeats within the polyphenolic protein can be rationalized as a physical feature that favors bonding to solids and can be explained in terms of thermodynamic and structural aspects of adhesive polymer theory [K.W. Allen, J. Adhesion 21, 261-277 (1987)] if it is assumed that the amino acid elements of the decapeptide control adhesion.

We have sought to study the mechanism of protein adhesion to metals by simulating the primary structure of the sea mussel polyphenolic protein as repeating poly(decapeptides) using various decapeptide precursors and by relating the adhesive quality of the synthetic polypeptide to the decapeptide sequence. This goal has required considerable effort to establish chemical and biological synthetic routes to polypeptides with internally repeating sequences so as to provide the biomaterial sources for experimentation. It has been a significant advance in this regard that we have established a

recombinant DNA system to express internally repetitive polypeptides from totally synthetic genes, and used this system to focus on technical barriers to utilization of microbial fermentations to produce repetitive polypeptides in an economic fashion. An important ancillary objective to our work has been to compare the enzymatic processing characteristics of several families of related peptides (most of them decapeptides) as a means to establish structure-function relationships that can guide future research devoted to the formation and processing of polypeptides with enhanced adhesiveness. This report covers the achievements in our contract efforts for the first two years of this program.

PEPTIDE SYNTHESIS

Oligopeptides were prepared on phenylacetamidomethyl (PAM) copolymer resins using an Applied Biosystems model 430A automated peptide synthesizer and symmetrical anhydride couplings with dicyclohexylcarbodiimide and t-BOC-protected amino acids. Syntheses were routinely performed in 0.5 mmol amounts with the carboxy-terminal amino acids already attached to the PAM resin. A list of the oligopeptides prepared under the support of this contract is shown in Table I. All of the amino acids were single-coupled except for GLUE-8 which required double-coupling for the addition of Phe to Lys-10. The efficiencies of the individual couplings were monitored by standard ninhydrin assays and, except for the Phe case just mentioned, were generally greater than 99%. Most PAM resins and protected L-amino acids were purchased from Applied Biosystems (Foster City, CA). Some of the mussel polyphenolic protein analogue peptides shown in Table I were prepared with certain sidechain groups blocked (e.g., ϵ-amino groups of Lys-blocked as the trifluoroacetamide derivative). Also, some of the analogue peptides were synthesized using protected 4-Hyp (4-hydroxy-L-proline) rather than proline; this was done to produce synthetic peptides more congruent in sequence and structure to decapeptide repeats found in the natural polyphenolic protein. The protected amino N-α-t-BOC-N-ϵ-TFA-Lys was purchased from Peninsula Laboratories (Belmont, CA) and 4-Hyp was purchased from Chemical Dynamics Corporation (South Plainfield, NJ). The synthesis of GLUE-4 and GLUE-5 required the use of N-α-t-BOC-N-ϵ-trifluoroacetyl-L-lysine-PAM resin. Since this material was not commercially available, it was synthesized in-house essentially by

Table I
GLUE DECAPEPTIDES

Glue-1	H_2N–ALA–LYS–PRO–SER–TYR–PRO–PRO–THR–TYR–LYS–OH
Glue-2	H_2N–ALA–LYS–PRO–SER–TYR–4 HYP–4 HYP– THR–TYR–LYS–OH
Glue-3	H_2N–ALA–LYS(TFA)–PRO–SER–TYR–4 HYP–4 HYP–THR–TYR–LYS–OH
Glue-4	H_2N–ALA–LYS–PRO–SER–TYR–4 HYP–4 HYP–THR–TYR–LYS(TFA)–OH
Glue-5	H_2N–ALA–LYS(TFA)–PRO–SER–TYR–4 HYP–4 HYP–THR–TYR–LYS(TFA)–OH
Glue-6	TFA–HN–ALA–LYS(TFA)–PRO–SER–TYR–4–HYP–4–HYP–THR–TYR–LYS(TFA)–OH
Glue-7	H_2N–ALA–LYS–PRO–SER–PHE–4 HYP–4 HYP–THR–TYR–LYS–OH
Glue-8	H_2N–ALA–LYS–PRO–SER–TYR–4 HYP–4 HYP–THR–PHE–LYS–OH
Glue-9	H_2N–ALA–LYS–PRO–SER–PHE–4 HYP–4 HYP–THR–PHE–LYS–OH
Glue-10	H_2N–ALA–LYS(TFA)–PRO–SER–PHE–4 HYP–4 HYP–THR–PHE–LYS(TFA)–OH
Glue-11*	H_2N–ALA–PRO–SER–TYR–4HYP–4HYP–THR–TYR–LYS–OH
Glue-12	H_2N–LYS(TFA)–PRO–SER–TYR–4HYP–4HYP–THR–TYR–LYS(TFA)–ALA–OH
Glue-13**	H_2N [ALA–LYS–PRO–SER–TYR–4HYP–4HYP–THR–TYR–LYS]$_2$OH
Glue-14**	H_2N [ALA–LYS–PRO–SER–TYR–PRO–PRO–THR–TYR–LYS]$_2$OH
Glue-15	TFA–HN–LYS(TFA)–PRO–SER–TYR–4HYP–4HYP–THR–TYR–LYS(TFA) ALA–OH

*nonapeptide
**eicosapeptide

following the published procedure of Tam et al [J. Amer. Chem. Soc. 108, 5242 (1986)]. All of the other peptide synthesis reagents and ninhydrin were from Applied Biosystems; dichloromethane (DCM) and dimethylformamide (DMF) were obtained from Burdick and Jackson (Muskegon, MI).

PEPTIDE CLEAVAGE AND PURIFICATION

The synthetic peptides were deprotected and cleaved from support resins using trifluoromethanesulfonic acid (TFMSA) rather than HF in most cases in order to reduce the hazards associated with the use of the more volatile HF reagent; all HF cleavage protocols proved effective with collagen and elastin peptide analogues, although some problem was noted with the conversion of methionine residues to sulfoxide or sulfonium salt forms. By contrast, initial attempts at one-step TFMSA cleavages with our GLUE compounds led to complex product mixtures resulting from side reactions and rearrangements among GLUE peptides. We have obtained HPLC and mass spectral data which conclusively demonstrate the need for an alternate peptide cleavage protocol for peptides related to repeating segments of the M. edulis polyphenolic protein. We therefore developed a two-stage TFMSA/TFA cleavage procedure which involves deprotecting the peptide under gentle conditions and washing away the newly released protecting groups before subjecting the deprotected peptide-resin to the more acidic conditions required for the cleavage of the peptide from the resin. The optimal conditions for the TFMSA deprotection stage of the two-step cleavage protocol was determined by ^{13}C NMR spectroscopy of peptide-resins swollen in DMF. Futhermore, monitoring the kinetics of the removal of chemical protecting groups by ^{13}C NMR appears to be a generally useful practice for other peptides prepared by solid phase means.

PEPTIDE PURIFICATION

Following cleavage, crude ether-precipitated peptides were passed over an AG1-X8 ion exchange resin (Bio-Rad) and were treated overnight with an excess of ammonium bicarbonate when appropriate to reverse N→O acyl shifts in Ser and Thr residues. HPLC analysis was routinely used both to verify complete reversal of the isomerization, and also to monitor the relative purity of the product before and after the bicarbonate treatment. The crude peptide was

then freeze-dried and purified by preparative HPLC using a Spectra-Physics 8100 system equipped with a 10 μm, 22 x 250 mm Vydac C-4 column. Sample components were eluted from the column using a binary gradient of 0-40% acetonitrile in 0.1% TFA in water at a flow rate of 3.5 mL/min. The major product peak was collected, freeze-dried and then subjected to a variety of analytical procedures to verify its identity and purity.

PEPTIDE CHARACTERIZATION

1. HPLC

HPLC-purified peptides were analyzed again by HPLC on a Spectra-Physics 8100 system used earlier. Sample components were monitored at a wavelength of 215 nm so the relative purity of the product could be assessed. Quantification of the product was also performed using standard curves constructed with bonafide materials when available. The final purity of all peptides as judged by HPLC was at least 95%; representative HPLC profiles for crude and purified preparations of GLUE-12 are shown in Figure 1. More recent applications of HPLC analysis include monitoring the removal of TFA-side chain protecting groups; an example of the kinetics of protecting group removal is shown in Figure 2.

2. Coupled HPLC/Diode-Array UV/Vis Spectroscopy

GLUES-2 through 4, 7 and 8, and 11 through 13 were analyzed using a Hewlett-Packard Model 1090M liquid chromatograph equipped with a built-in diode-array detector. HPLC and detector parameters and evaluation of chromatograms and spectra were carried out with a Hewlett-Packard HPLC ChemStation supported by a HP900 series 300 computer. Chromatograms and spectra were plotted using a HP7440A eight-pen plotter. Chromatographic separations were performed with the analytical C-4 Vydac column as described above, at a column temperature of 40^{o}C.

3. Mass Spectrometry

Fast atom bombardment-mass spectrometry (FAB-MS) was performed using a VG analytical ZAB-MF double focussing mass spectrometer. Peptide samples were

dissolved in glycerol and were bombarded by Xenon atoms at 8 Kev. Sputtered ions were recorded using an ion energy of 8 Kev and magnetic scanning in the range 10-3000 mass units with a VG 12-250 data system. Peptide sequence data was extracted from the aggregate information provided by the positive and negative ion spectra. In many cases, complete and unambiguous amino acid sequences beginning at either the amino or carboxyl terminus could be assigned from the FAB-MS data. A representative positive ion mass spectrum for purified decapeptide GLUE-12 is shown in Figure 3; the intense mass lines from this spectrum, characteristic of a highly purified material, were used to construct sequence information based upon the mass series A, B, C" and Y" (cf. Table II). The complete amino acid sequence of GLUE-12 could be derived from the A_n series and confirmed by the overlapping regions of the B_n and $C"_n$ series.

4. HPLC-Thermospray Mass Spectroscopy

Purity analysis was performed on GLUES-7, 8, 11, 12 and 13. Spectra were obtained using a Varian model 5000 HPLC unit connected to a VG Masslab 30-250T triple quadrupole mass spectrometer. Post-column addition of buffers was achieved with a Milton Roy Constametric IIG pump and the total effluent was passed through a Milton Roy Spectromonitor III variable wavelength UV detector. Thermospray mass spectra were recorded using the first ion detector at the exit from Q1 with a mass range of 10-3000 units. Data acquisition and instrument control were done with the VG 12-250 data system.

5. Amino Acid Analyses

Amino acid composition of many purified peptides was determined by OCS Laboratories, Inc., Sanger, TX and in all cases was found to be consistent with the expected amino acid content.

6. NMR Spectroscopy

^{13}C, ^{19}F and ^{1}H NMR spectra were obtained on a Varian XL-400 spectrometer at Allied-Signal; in the case of GLUE-13, one- and two-dimensional NMR spectra were also collected at 490 or 500 MHz in the laboratory of Dr. James

Table II

Sequence ions identified from FAB-MS analysis of GLUE-12

	Sequence	Mass	A	B	C"	Y"	
	H	1					
1	Lys TFA	224	197	225			10
2	Pro	97	294	322	339	1149	9
3	Ser	87	381	409	426		8
4	Tyr	163	544	572		967	7
5	Hyp	113	657	685	702	804	6
6	Hyp	113	770		815	691	5
7	Thr	101	871	899	916		4
8	Tyr	163	1034	1062	1079	477	3
9	Lys TFA	224	1258		1303	314	2
10	Ala	71	1329	1357			1
	OH	17					

Prestegard, Department of Chemistry, Yale University. The spectra recorded at Allied-Signal were obtained using a double precision data acquisition with a 45^{o} flip angle and a pulse recycle time of 0.6 seconds. Typically, spectral data were collected in overnight runs. NMR spectra were not only used to determine sample purity (primarily ^{1}H and ^{13}C NMR) but also to monitor the course of removal of side chain protecting groups (^{19}F NMR) and to examine solution conformation characteristics (^{1}H NMR) (Figures 4 and 5).

PEPTIDE POLYMERIZATION AND ADHESIVES TESTING

A large number of methods are known for synthesizing polypeptides by the direct polycondensation of monomers or oligopeptides under mild conditions. We have developed several strategies for chemical polymerization of synthetic peptides based upon the condensation properties of one coupling reagent, diphenylphosphoryl azide [DPPA, $(C_6H_5O)_2P(O)N_3$]. DPPA is a stable, non-explosive liquid (bp $157^{o}C$) and was originally reported in 1972 as a convenient reagent for inducing formation of amide bonds by a modified Curtius reaction [J. Amer. Chem. Soc. 94, 6203 (1972); Int. J. Biol. Macromol. 2, 53 (1980)]. The amide bonds are preferentially formed in the presence of a strong organic base such as triethylamine and in aprotic solvents such as dimethylsulfoxide (DMSO). The advantages of DPPA in promoting amidation reactions have been shown to include racemic-free condensations of amino acids of peptides and minimal side reactions with -OH groups when they are present in unprotected peptide side chains. DPPA also offers analytical advantages since peptide polymerizations with DPPA at room temperature proceed slowly enough to allow polymerization kinetics to be carefully analyzed.

Preparation of several polypeptides as well as poly(decapeptides) with varying levels of protection of the Lys ϵ-amino groups was achieved in our laboratory with DPPA as the polymerization agent. In some cases, polymerization reaction conditions such as time, temperature, reactant concentrations, coreactants and choice of organic base cosolvent were varied to study their impact on the course of peptide polymerization.

We initially polymerized the dipeptide L-Ala-Gly with DPPA in DMSO (Table III). The reaction was stopped after 48 hours and the polymeric product was purified. The pure polypeptide was analyzed by IR (as a KBr pellet) and intrinsic viscosity and light scattering methods were employed to determine

Table III

PEPTIDES AND AMINO ACIDS POLYMERIZED USING DPPA WITH DMSO AS SOLVENT

Compound	MW Range Achieved
L-Ala-Gly	9,000-44,000
L-Val-Pro-Gly-Val-Gly	18,000-25,000
GLUE-2	7,000
GLUE-12	26,000-50,000
L-Dopa	4,000-5,500

the molecular weight. The intrinsic viscosity in dichloroacetic acid at 25°C was determined to be η=0.14 dl/g. Literature values for the calibration constants k and a in the Mark-Houwink equation $\eta = k M^a$ were taken for poly(N-benzyl-L-Ala) and poly(γ-benzyl-DL-Glu) to obtain a number average molecular weight (M_n) range of 9,000-14,000. However, light scattering measurements provided a value for M_w (weight average molecular weight) of >40,000; assuming a typical Gaussian relationship between M_w and M_n of $M_w = 2M_n$, a discrepancy between these two measurements exits.

The IR spectra for poly(L-Ala-Gly) support peptide polymerization. Amide peaks characteristic of -NH and C=O absorption bands appeared at 3300 and 1640 cm^{-1} (Figure 6). Another strong N-H absorption was seen at 1525 cm^{-1} with a weaker harmonic at 3050 cm^{-1}. The relative intensities of the C-H absorption at 2980 cm^{-1} and N-H absorption at 3300 cm^{-1} varied as the dipeptide monomers were converted to the polymers (Figure 6). Similarly, ratios of the amide peaks at 1640 cm^{-1} and 1525 cm^{-1} varied in accordance with the appearance of the poly(dipeptide). A small absorption peak at 2130 cm^{-1} is attributed to the acyl azide at the C-terminal of the polymer chain (i.e., from the N_3 group of DPPA). This suggests that a fraction of the reaction might have proceeded via an azide coupling.

Poly(L-Val-Pro-Gly-Val-Gly) was also prepared as an elastin analogue with DPPA as the polymerization reagent (Table III). After 9 days of reaction, polymer product was precipitated and dialyzed against water for 48 hours using 8-10 kDa (kilodaltons) cut-off dialysis tubing. Non-dialyzable polypeptides were recovered in 30% yield and molecular weight was determined by intrinsic viscosity (η=0.22 dl/g) and gel permeation chromatography. Characteristic polymer peaks were also seen in NMR and IR spectra.

One set of synthetic polymerization reactions with the elastin analogue pentapeptide was carried out with pyridine or triethylamine as cosolvent bases. Reactions were allowed to run for 17 days at room temperature. The individual reaction mixtures were then diluted separately with deionized water and dialyzed for 40 hours as described above. No solids were recovered in the vessel containing reaction material prepared with pyridine. However, the polymer yield was greater than 50% when triethylamine was used. A stronger base is therefore more effective in DPPA-mediated polymerizations. We determined the molecular weight of poly(pentapeptide) in this batch by estimating the amount of free amines in both the polypentapeptide and the

pentamer by quantitative ninhydrin assays. The M_w calculated from the ratio of absorbance values at 570 nm for polymer and pentamer was greater than 25,000.

The GLUE decapeptides GLUE-2 and GLUE-12 were polymerized in a similar fashion. Since neither of the Lys residues in GLUE-2 was blocked, we anticipated extensive branching during DPPA-induced peptide polymerization. Polymerization of both decapeptides for one week led to the isolation of off-white precipitates recovered in greater than 60% yield. Ninhydrin assays on the polypeptide products gave polymerization numbers of 7 for poly(GLUE-2) and 20 for poly(GLUE-12). The lower degree of polymerization for poly(GLUE-2) can be ascribed to the build-up of positive charge on the growing poly(GLUE-2) polymer units which at some point inhibits effective unit condensations because of charge repulsion. Charge build-up is not anticipated for Lys-protected GLUE-12.

X-ray scattering data with blocked poly(GLUE-12) has given a radius of gyration R_g=30 Å with limited heterogeneity in polypeptide size. Blocked poly(GLUE-12) gave excellent scattering profiles due to the numerous pendant electron dense fluorine atoms per molecule. The large value of R_g suggests a high molecular weight polypeptide (>50 kDa) or a polypeptide of lesser mean molecular weight with a more open conformation porous to solvent; the latter possibility seems more likely since ninhydrin assays on this polypeptide preparation gave a molecular weight of 26 kDa.

L-Dopa has been polymerized with DPPA to obtain a low molecular weight (4.0-5.5 kDa) polymer. Confirmatory evidence for successful polymerization was obtained by ninhydrin assay, NMR and IR spectroscopy.

Although analytical methods other than ninhydrin assays, (intrinsic viscosity measurements, gel permeation chromatography, NMR, IR, fast protein liquid chromatography (FPLC), light scattering and small x-ray scattering methods) have been screened for the determinations of molecular weights of polydecapeptides, these methods suffer from a number of drawbacks, such as unproven assumptions about polypeptide conformation or lack of appropriate calibration constants. Thus, ninhydrin assays to determine the number of free primary amino groups appear to provide the most consistent endpoints in molecular weight determinations. A summary of mean molecular weights for various polypeptides prepared by the DPPA method is listed in Table IV.

Table IV

MEAN MOLECULAR WEIGHTS DETERMINED FOR POLYPEPTIDES PREPARED WITH DPPA

Polypeptides	Solvent	Analytical Medthod	Molecular Weight
Poly(L-Ala-Gly)	Dichloroacetic Acid (DCA)	Intrinsic Viscosity	8,700 (η = 0.14)
Poly(L-Val-Pro-Gly-Val-Gly)	DCA	Light scattering	44,000
	DCA	Intrinsic viscosity	22,000 (η = 0.22)
	Water	Gel permeation	18 – 25,000
	Water	FPLC	16 – 18,000
	Ninhydrin Reagents*	Ninhydrin	16,500
Poly (Glue-12)			
– (Blocked Lysines)			
6-Days Polym.	Ninhydrin Reagents	Ninhydrin Assay	26,000
11-Days Polym.	Ninhydrin Reagents	Ninhydrin Assay	35,000
4-Weeks Polym.	Ninhydrin Reagents	A. Ninhydrin Assay	40,000
	DMSO	B. Intrinsic Viscosity	(η = 0.76)
	DMSO	C. Light Scattering	27,000
7-Weeks Polym.	Ninhydrin Reagents	Ninhydrin Assay	50,000
– (Deblocked Lysines)			
11-Days Poymer.	6M-Guanidine-HCl in PBS	Pharmacia FPLC	25,000
Poly(glue-2)			
11-Days Polym	Ninhydrin Reagents	Ninhydrin Assay	7,000

* Solvents: Ethanol, Phenol, Pyridine

Polymeric GLUE-12 was deblocked by treating the polypeptide with methanolic piperidine solution (methanol:piperidine:water, 7:2:1). This rendered the polymeric product water soluble as desired for adhesive testing. Deblocked materials were desalted prior to further manipulation by eluting the samples from a Sephadex G-25 column with water as the mobile phase. These polypeptides were used in adhesive tests directly following lyophilization and resuspension in water.

Tests were performed on a variety of polypeptides to determine the lap shear strengths of adhesion by glueing together polished aluminum plates. The lap shear joints were tested according to ASTM test procedure D1002-72 using recommended 1/2" overlaps. Mechanical tests were carried out on an Instron instrument using typically 500 psi full scale load and 1/5"/min. head speed. The aluminum test plates were fabricated from 1/16" stock with external dimensions of 4" x 1".

In most experiments, the polypeptide adhesive bond was prepared by dissolving preweighed amounts of polypeptide in water and pouring a known amount of the solution on clean 1/2" square areas of the aluminum surfaces with or without various coadditives. Batches of at least three pairs of plates were used in the measurements of average adhesive strength. Satisfactory lap shear strengths have been obtained for the sea mussel polyphenolic protein analogue poly(GLUE-12) treated with the enzyme mushroom tyrosinase (Table V). Without the enzyme treatment, the lap shear strength of poly(GLUE-12) was cut by half. Addition of the mushroom tyrosinase to the poly(GLUE-12) substrates in situ substantially increased the degree of successful bonding to the aluminum test plates so as to provide a degree of adhesiveness that was similar to that found with the poly(Lys) polypeptides tested in parallel runs. No adhesive strength could be measured for poly(GLUE-2), MW 7,000, under standard conditions because of its inability to solubilize in water. Partially dissolved poly(GLUE-2) in methanol displayed almost no tackiness or adhesion when tested.

CONFORMATIONAL ANALYSIS OF THE EICOSAPEPTIDE GLUE-13

We have completed circular dichroism and ^{1}H nuclear magnetic resonance (NMR) studies of GLUE-13 in solution. As described in our contract renewal proposal, the one dimensional and two dimensional ^{1}H NMR spectra do not

Table V

LAP JOINT SHEAR STRENGTH OF POLYPEPTIDES ON ALUMINUM IN WATER

Polypeptide	Conc. in Water (mg/45 μL)	Lap Shear Strength* 1/16" Plates (kg/cm^2)
Poly(glue-12) (Blocked Lysines)	Water Insol.	--
Poly(glue-12) M.W.>40,000 (Deblocked and Enzyme Treated)	5	5.6
Poly (Lys) HBr M.W.>70,000	5	5.7
Poly (Lys) HCl M.W.>40,000	5	7.4
Devcon 2-Ton Clear Epoxy (Thin Coats)	–	17.5

* Average of 3 Tests

support a regular structure for this eicosapeptide in solution, suggesting that GLUE-13 has a relatively open conformation or set of conformations in water. New CD data were generated in cooperation with Dr. Ken Breslauer, Department of Chemistry, Rutgers University. CD profiles were collected in two separate experiments over the temperature ranges 5-55°C or 5-65°C; the data were deconvoluted with the computer program PROSEC. The deconvoluted data suggest a signal that is 60-65% random coil, 7-16% beta turn, and 20-45% beta strand with no alpha helix. The original spectra themselves (Figure 7) appear consistent with at least one regular conformation, a tight left-handed helix with 2.5 residues per turn (S. Krimm, personal communication). Such regularity conflicts with the results obtained by ^{1}H NMR and leaves room for further experimentation. One explanation for the difference would be that the eicosapeptide GLUE-13 has several energetically favored conformations, all of which exhibit some degree of open or fluid structure so as to render any NMR results inconclusive. We are excited that a significant fraction of non-random structure is seen by CD with GLUE-13 and we plan to extend our CD studies to the decapeptide homologue GLUE-2 to see if this shorter peptide also has a tendency towards short range order. Our observations of potential beta strand-beta turn structures within GLUE-13 solution could provide a physical basis for the observed differences in this rate of enzymatic oxidation and provide clues for the nature of crosslinks between lysine and tyrosine residues within decapeptide repeats.

SYNTHESIS OF POLYPEPTIDE-CONTAINING BLOCK COPOLYMERS

We have demonstrated a route for the production of a class of polypeptide-containing block copolymers during this contract in which macroscopic properties of the polypeptides are modified by the presence of other organic polymer block segments. The process involves polymerization of the decapeptide (GLUE-12) with blocked lysines with DPPA for 24 hours. A parallel reaction was carried out to polymerize ϵ-aminocaproic acid with DPPA for the same period of time. Portions of each sample were then mixed together and the polymerization was continued for the mixture and the parent compounds for 11 days. Polymeric products from each reaction were purified in greater than 60% yield. They were dried *in vacuo* for two days and were analyzed to determine the thermal transition temperatures by thermogravimetric analysis (TGA) and

differential scanning calorimetry (DSC). A Dupont 9900 thermal analyzer with a DSC cell in an argon atmosphere was used for the analyses. The polymer samples (each about 15 mg) were crimped in an aluminum pan and heated at $10^{o}C$/min. After initial heat-up, the samples were held at $250^{o}C$ for 5 minutes prior to either programmed cooling ($10^{o}C$/min) or quenching in liquid nitrogen. A subsequent reheat was carried out under the same conditions as the initial heat-up. DSC data (Figure 8) revealed that the polydecapeptide poly(GLUE-12) did not show any definite glass temperature (T_g), crystallization temperature (T_{ch}) or melting temperature (T_m). By contrast, the block copolymer of Nylon-poly(GLUE-12) displayed a distinct DSC temperature profile with T_g=$34^{o}C$, T_{ch}=$108^{o}C$ and T_m=$174^{o}C$. This profile was similar in pattern to those observed for the Nylon polymer (Figure 8), thus suggesting that the processing temperature for a block copolymer of this type would be dictated by the nature and content of the organic polymer blocks. The nylon content of the block copolymer used in Figure 8 was estimated by GC method to be about 65%. These results have some obvious implications for the processing of high molecular weight synthetic or natural polypeptides with bioadhesive properties.

ENZYMATIC OXIDATION OF GLUE PEPTIDES WITH MUSHROOM TYROSINASE

1. Spectrophotometric Studies

L-Lys and mushroom tyrosinase (E.C. 1.14.18.1) were purchased from Sigma Chemical Co. and used without further purification. Oxidation studies were normally performed with 50 μg of the tyrosinase, corresponding to about 110 units based on the enzyme specific activity (2260 units/mg solid) supplied by the manufacturer. Due to nonhomogeneity of freeze-dried preparations, the specific absorbance by the enzyme was adjusted to a standard value of 0.15 A_{280} units/mL buffer before use. The GLUE-2, -3, -4, -7, -8 and GLUE-12 peptides (Table I) were separatly mixed with the enzyme and the unfractionated oxidation reaction mixtures were studied spectrophotometrically to determine gross reaction changes. Reaction mixtures were scanned over the 190-600 nm wavelength region using a Perkin-Elmer model 553 Fast Scan UV/vis spectrophotometer. Spectra were recorded with a Perkin-Elmer R100A recorder at 1 AUFS at a scan rate of 240 or 480 nm/min. Reactions were generally

initiated by the addition of substrate at a concentration of 0.075 mM in a final volume of 1 mL of 0.1 M sodium phosphate buffer, pH 7. In some experiments, a five- or ten-fold molar excess of L-Lys relative to peptide substrate concentration was added to the reaction mixture. Control experiments to measure the effect of L-Lys on the ability of mushroom tyrosinase to oxidize L-Tyr were also carried out. After initiating the reaction, spectra were recorded at timed intervals and air at room temperature was bubbled through the sample cuvette between scans or at 5 min. intervals throughout the duration of the experiment.

The results of these oxidation studies with several of the synthetic decapeptides are shown in Table VI. The GLUE peptide substrates absorb primarily in the UV with a maximum of about 275 nm and with a shoulder above 280 nm. Upon exposure to tyrosinase, the absorption maximum shifted to about 280 nm, due to conversion of tyrosine residues to Dopa, with a shoulder at about 310 nm. These spectral features were observed for all the GLUE compounds examined thus far. The increases in 280 nm absorbance as a function of time (dA_{280}/dt) for GLUE-2, -3, -4 or GLUE-12 (the latter in a blocked or unblocked LYS state) are listed in Table VI.

Our interpretation of these data rests on the premises that: (1) cross-linking involving the ϵ-amino group of Lys residues occurs primarily with oxidation products of the tyrosine residues, i.e., cross-linking can take place only after oxidation of the tyrosines to Dopa or its quinone derivatives; (2) as the Dopa or quinone residues participate in crosslinking reactions, the characteristic absorbances in the 260-280 nm range are reduced and UV absorbance is shifted to longer wavelengths and is spread over a wider range of wavelengths; and (3) crosslinking reactions induced by oxidative transformations of the GLUE peptide substrates can be intra- or intermolecular in nature. The simplest oxidation results were those for GLUE-12, where both Lys ϵ-amino groups are blocked and neither Lys residue can be expected to participate in crosslinking reactions, leading to rapid development of a characteristic Dopa/quinone spectrum. As anticipated, the dA_{280}/dt is the highest (Table VI). The addition of exogenous Lys to GLUE-12 oxidation reaction mixtures had minimal effect. However, a ten-fold excess caused the rate to be suppressed, presumably by covalent reactions of the exogenous Lys with oxidized forms of tyrosines. The GLUE-2, -3, and GLUE-4 compounds, each having at least one unblocked Lys residue, would be expected to form intra- or

Table VI

Rate of development of 280 nm absorbance for GLUE decapeptides exposed to mushroom tyrosinase

Substrate	$dA_{280}/dt \times 10^{-5}$ (in sec^{-1})
GLUE-2	3.4
GLUE-2 + 5-fold molar excess of free Lys	3.3
GLUE-2 + 10-fold molar excess of free Lys	3.7
GLUE-3	3.7
GLUE-3 + 5-fold molar excess of free Lys	3.8
GLUE-3 + 10-fold molar excess of free Lys	3.6
GLUE-4	5.8
GLUE-4 + 5-fold molar excess of free Lys	3.7
GLUE-4 + 10-fold molar excess of free Lys	4.0
GLUE-12	17.7
GLUE-12 + 5-fold molar excess of free Lys	16.4
GLUE-12 + 10-fold molar excess of free Lys	9.8
GLUE-12, deblocked	11.1
GLUE-12, deblocked, + 5-fold molar excess of free Lys	9.6
GLUE-12, deblocked, + 10-fold molar excess of free Lys	9.2

interchain crosslinks. A much reduced value for the dA_{280}/dt was observed, as was anticipated. The oxidation patterns shown in Table VI are also consistent with the supposition that the carboxyl terminal lysines are more active in cross-linking than amino terminal lysines and certainly far more active than exogenous free Lys. Support for this contention is provided by the apparently minimal effect of adding free Lys to GLUE-2 or GLUE-3, a response which can be explained either in terms of more rapid intramolecular crosslinking of oxidized Tyr-9 intermediates with adjacent Lys-10 residues, or by rapid intermolecular crosslinking between the Lys residues on one molecule and the oxidized Tyr residues on another molecule. Additional information in this regard was obtained by coupled HPLC-UV/vis spectroscopy studies on GLUE peptide oxidation mixtures as a function of time (see below). Since intramolecular reactions generally proceed at a much more rapid rate in the absence of molecular steric hindrance than intermolecular reactions, exogenously supplied Lys might be expected to have little chance to intervene and react with GLUE-2 or GLUE-3, and possibly with GLUE-4. GLUE-4 presents a different situation since Lys-10 is blocked as a trifluoracetamide and cannot form Tyr-9/Lys-10 intramolecular crosslinks. From the data in Table VI, it is apparent that oxidized GLUE-4 populations have a significant portion of material that is stabilized as Dopa/quinone species. Addition of free Lys to GLUE-4 reaction mixtures leads to reactions between the free Lys and Dopa/quinone species and the value of dA_{280}/dt is similar to that seen with GLUE-2 and GLUE-3. Given the availability of LYS-2 for intramolecular reactions in GLUE-4, we then interpret the difference in reactivity of GLUE-4 when compared to GLUE-2 or GLUE-3 to the decreased availability of LYS-2 for intra- or intermolecular crossslinking. Despite this difference, there are several routes for intramolecular crosslinking in this decapeptide, e.g., between Lys-2 and either Tyr-5 ot Tyr-9. The possibility of a crosslink via the α-amino group of Ala-1 also cannot be ignored. An intramolecular crossslink to Tyr-9 in GLUE-4 thus presents obvious constraints on the final peptide conformation, requiring the juxtapositioning of distal components in this decapeptide. It should also be pointed out that another possible route of intramolecular crosslinking, Lys-2 to Tyr-5, cannot be ruled out. However, our observations based upon oxidation results with GLUE-7 and GLUE-8 (see below) makes this possibility less likely.

The enzymatic oxdation data collected with TFA-blocked or deblocked GLUE-12 stand apart from the other results and present an interesting case for discussion. In addition to the rapid build-up of material absorbing at 280 nm during the oxidation of blocked GLUE-12 (Table VI), it is now apparent that there was also an appearance of a 340 nm component at later time points in the course of blocked GLUE-12 oxidation reactions. No such component was observed with GLUE-2, -3 or -4. According to Linder and Dooley [Proc. 3rd Int. Biodegration Symp., 465-494, (1976)], a 340-350 nm absorption band is displayed by quinone crosslinked proteins. We have therefore concluded that some intermolecular or intramolecular crosslinking involving functional groups other than ϵ-amino groups probably does occur with GLUE-12. A component with absorbance at 348 nm was also observed by coupled HPLC-UV/vis spectroscopic analysis of GLUE-12 oxidation reaction products and is discussed in greater detail later on. The dA_{280}/dt value for deblocked GLUE-12 was significantly less than that for blocked GLUE-12 (Table VI), indicating a greater initial rate of crosslinking, but this value was more than seen with GLUE-2 (which presumably is a homologue to deblocked GLUE-12). The reason for this discrepancy is not clear, although we can rule out incomplete deblocking of GLUE-12 as a contributing factor since the removal of pendant TFA moieties in GLUE-12 was atleast 98% complete (cf. Figure 2). Deblocked GLUE-12 is a slightly permuted sequence of GLUE-2 and this permutation may be the important factor in our results since Lys-10 in GLUE-2 and GLUE-3 is considered to be an essential crosslinking agent. This may be due to the placement of Lys-10 at the carboxyl terminus of these decapeptides, while the Lys residue proximal to the carboxyl terminus in deblocked GLUE-12 is actually in the penultimate position. The amino acid sequence difference between GLUE-2 and deblocked GLUE-12 appears to have a substantial influence on the enzymatic oxidation rates of these peptides. Thus, the addition of exogenous Lys to deblocked GLUE-12 oxidation reaction mixtures has only a limited influence on the appearance of material absorbing at 280 nm with time, much like the results obtained with GLUE-2 (Table VI). Neglecting charge effects that might be associated with using exogenous free Lys, these results indicate that the rate-limiting reaction step in the oxidation of deblocked GLUE-12, in comparison to GLUE-2, is associated with the oxidation of tyrosine residues. It appears valid at this point in our work to accept that charge effects of free Lys are not important. We therefore have concluded that the difference

in oxidation rate for the two peptides (GLUES-2 and deblocked GLUE-12) might be attributable to different peptide secondary structures induced by altered decapeptide sequences. In other words, the more rapid change in A_{280} in oxidation mixtures of deblocked GLUE-12 compared to GLUE-2, -3 or GLUE-4 might merely reflect a preference of the mushroom enzyme for tyrosyl residues at position #8 over position #9. Further insight into the relative importance of decapeptide amino acid sequence influence on oxidation behavior can be expected from the oxidation studies with blocked and deblocked GLUE-5 (Table I). It must be noted that our oxidation studies to date were performed using decapeptides in fairly dilute solution (0.075 mM), thereby favoring intra- over intermolecular reactions, and therefore might not be directly comparable to reactions performed at much higher concentration or to oxidations of the natural sea mussel polyphenolic proteins.

2. Chromatographic and Mass Spectral Studies

The decapeptides GLUE-2, -3, -4, -7, and -8 and GLUE-12 were oxidized with mushroom tyrosinase in phosphate buffer, pH 8.0, in the presence or absence of ascorbic acid. Ascorbic acid was used in these experiments in order to retard oxidation of tyrosines beyond the Dopa state. The products of these reactions were analyzed as a function of time by HPLC, coupled HPLC-UV/vis spectroscopy, and, in certain instances, by coupled HPLC-mass spectrometry.

HPLC analysis of GLUE decapeptide reaction mixtures incubated with an excess of ascorbic acid showed mostly unreacted GLUE decapeptide along with small amounts of its Dopa derivatives (Figure 9). Adequate oxidation of Tyr in GLUE-12 to Dopa residues was observed when the oxidation experiments were performed using, for example 0.05 mM GLUE-12 in the presence of 5-20 nM ascorbic acid for a period ranging upto three hours. No additional peaks characteristic of other oxidation products were observed under such controlled conditions. However, after 18 hours, new peaks eluting at 12-13 min. (Figure 10) were observed. These peaks were identified by coupled HPLC-MS to be deblocked GLUE-12 hydrolysis products. These products rarely amounted to more than 2% of the total. The use of suboptimal concentrations of ascorbic acid in some instances led to the oxidation of the Dopa moiety to the quinone derivatives after long incubation periods.

The preferential oxidation of Tyr-9 over Tyr-5 was evident when the HPLC profiles of the oxidized decapeptides GLUE-2, -7 and -8 were compared. The rate of tyrosine oxidation in GLUE-7 (with Phe at position #5) was comparable to that seen with GLUE-2 in presence or the absence of asorbic acid as judged both by HPLC and by UV/vis spectroscopic profiles (Figure 11A). By contrast, GLUE-8 (with Phe at position #9) underwent almost no oxidation as judged by HPLC under identical reaction conditions (Figure 11B). Additional supporting evidence for positional influences on the oxidation of Tyr residues came from the oxidation of GLUE-12 in the presence of ascorbic acid, where the major product identified by mass spectrometry was a mono-Dopa decapeptide. One of the minor products that appeared after 18 hours of enzyme treatment was identified by mass spectrometry to be a di-Dopa decapeptide (Figure 10). These data taken together argue that mushroom tyrosinase preferentially oxidized the Tyr residues at position #9 of GLUE-2 or GLUE-7 (and position #8 of GLUE-12) over those at position #5 (or position #4 for GLUE-12). This result is the same reported by Marumo and Waite [Biochim. et Biophys. Acta 872, 98-103 (1986)] for the oxidation of a decapeptide similar to GLUE-1 (Table 1).

Evidence for the involvement of Lys ϵ-amino groups in crosslinking reactions was found by comparison of the HPLC profile and UV/vis spectroscopic data for oxidized GLUE-12 in the absence of ascorbic acid (Figure 12) to the profiles for oxidation of GLUE-2, -3, and -7 under similar conditions (Figures 9, 11). Without ascorbic acid, the oxidation products with blocked GLUE-12 form sharp peaks. However, small (<2%) quantities of the same deblocked/ hydrolysed products observed earlier when GLUE-12 was incubated with ascorbic acid were also formed in the absence of ascorbic acid. The fact that some degree of unblocking of Lys residues in GLUE-12 occurs during enzymatic treatment explains why we may have intramolecular and/or intermolecular crosslinking with oxidized GLUE-12. A unique reaction component was seen during oxidation of GLUE-12 with an elution time of about 19.6 min. This component was present in small amounts after 55 min. of GLUE-12 oxidation, in major amounts after 105 min. and was the predominant species after 180 min. of reaction (Figure 12). Examination of the UV spectrum for this component and the other major components observed during GLUE-12 oxidation (Figure 13) revealed that this component has spectral peaks at about 270 and 348 nm. The nature of this material was partially determined by performing coupled HPLC-thermospray mass spectrometry on all the major oxidized products of

GLUE-12 (Table VII). The 19.6 min. elution peak material has a mass spectrum consistent with the formation of intramolecularly crosslinked monoquinone (with a mass difference of +12). The reason why this material is the predominant product while the early eluting GLUE-12 hydrolysis peaks are minor products is best explained by the hydrolysis step being rate limiting, with intramolecular reactions occurring as rapidly as deblocked GLUE-12 decapeptides are generated.

When incubated overnight in the absence of ascorbic acid, GLUE-12 formed products which appeared as a broad peak (retention time 20-28 min.) Coupled HPLC-thermospray mass spectrometric analysis (Figure 14) of this broad peak indicated a molecular weight consistent with an intermolecular GLUE-12 dimer (Table VII). However, the resolution was not sufficient to allow speculation regarding the nature of the intermolecular crosslink(s). Semipreparative HPLC is currently being used to collect enough of this material for FAB-MS analysis in order to verify the dimerization and to possibly define the mechanism involved in its formation.

A chromatographic and mass spectral analysis of the oxidation products of the eicosapeptide GLUE-13 was initiated recently. The results will be forthcoming.

SYNTHETIC GENE CASSETTE AND COCASSETTE CLONING

The construction and expression of synthetic gene cassettes for polypeptide block polymers and/or copolymers with internally repeating amino acid sequences offers a unique route to synthesizing analogues of structural proteins with interesting physical properties. We chose to express synthetic genes in the gram-negative bacterium Escherichia coli because of the wealth of genetic information on gene regulation and gene expression available for this microorganism. The application of recombinant DNA methods for constructing and establishing gene expression systems in E. coli offers a viable and economic alternative to chemical synthesis as a source of homogeneous preparations of polypeptides. Unlike chemically polymerized peptides which will exhibit a distribution of sizes, polypeptides expressed from genes or gene cassettes on an engineered plasmid expression vector should be uniform in length. Since variable polymer chain length would no longer be a problem, biologically synthesized materials would be preferable for controlled modification by chemical or enzymatic methods.

Table VII

Components found in GLUE-12 oxidation incubation mixtures by coupled-HPLC/thermospray MS

Retention Time	Mass Difference	Comments
14.9 min.	+32	Probably di-Dopa derivative. Seen best in runs containing ascorbic acid.
15.6 min.	+16	Mono-Dopa derivative, major product in runs containing ascorbic acid.
16.3 min.	+0	Parent GLUE-12 substrate
18.9 min.	+30	Could be mon-dopa + mono-quinone
19.6 min.	+12	Has molecular weight consistent with intramolecularly crosslinked quinone monomer. Major product after 180 min. in the absence of ascorbic acid. Has the unique 270/348 nm absorbance maxima.
20-28 min.	ca. dimer	Broad, poorly resolved peak. Seen only after long incubation periods. The thermospray-mass spectra spectrum is shown in Figure 14.

1. Vector Construction

We have prepared several specialized gene expression vectors containing specific cloning sites, strong and regulatable bacteriophage promoters upstream of the multiple cloning sites, and one or more antibiotic selection markers. All of the constructed vectors are based on the expression vector pJL6 (Table VIII) developed in the NIH laboratories of Drs. Don Court and Takis Papas. The transcriptional initiation region of pJL6 is based upon the p_L promoter of bacteriophage lambda. The translational initiation and termination regions and ClaI cloning site of pJL6 can be conveniently removed as a small NdeI-HindIII fragment. Derivatives of pJL6 can then easily be generated by cassette mutagenesis to produce a family of related vectors useful for the production of various polypeptide polymers or block copolymers (based upon contiguous gene cocassettes).

A family of expression vectors has been generated for the construction and expression of analogue peptides for collagen, the M. edulis polyphenolic protein, and elastin. The basic scheme by which these vectors were made relies upon initial modification of pJL6 to generate the progenitor plasmid pAV02 (Table VIII) followed by cassette mutagenesis of the small NdeI-HindIII region of pAV02. pAV02 contains a synthetic consensus T7 RNA polymerase promoter located on the upstream side of the cloning site (i.e., the DNA region proximal to the p_L promoter), while a synthetic SP6 promoter has been inserted on the downstream side. Transcription through the cloning site from either promoter allows for in vivo or in vitro characterization of either strand of any DNA insert. The T7 promoter also provides an alternative for gene cassette expression when used in conjunction with a cloned gene for T7 RNA polymerase. The important properties of the plasmids pAV1-pAV8 derived from pAV02 are summarized in Table IX. The rop gene controls plasmid copy number and was deleted during construction of pAV01 and pAV02; as a result, the copy number of pAV02 derivatives is about 3-5 fold higher than pJL6. Plasmids pAV2, pAV4, pAV6 and pAV7 are particularly valuable since the enzymes used to cleave the cloning region create asymmetric half-sites which impart directionality upon gene cassettes during cloning. This facilitates the cloning of multiple cassettes and cocassettes by eliminating the need to screen recombinant plasmids for the appropriate orientation of cloned inserts.

Table VIII. Strain and Plasmid List

Designation	Genotype	Source
	Parent Strains	
DC1138	r⁻m⁻, pro, leu, Δ(srlR-recA)306::Tn10, (λ)	Tom Patterson
DC1139A	r⁻m⁻, pro, leu, Δ(srlR-recA)306::Tn10, [λdef, ΔBam, cI857(Ts), ΔH1]	Tom Patterson
JM109	hsd R17 (rk⁻, mk⁺), rec A1, end A1, gyr A96, thi, sup E44, rel A1, Δ(lac-pro AB), (F tra D36, pro AB, lacI^{q}Z ΔM15)	Stratagene
TAP129	lac Z (Am), trp (Am), pho (Am), sup C (Ts), mal, rpsL, phe, rel, [λdef, ΔBam, rex::Km⁺ cI857 (Ts), Δ (cro-bioB)]	Tom Patterson
TAP130	TAP129, hptR165	Tom Patterson
IG109	TAP129, Δ(srlR-recA)306::Tn10	in house
IG110	TAP130, Δ(srlR-recA)306::Tn10	in house
	Parent Plasmids	
pJL6	bla	Don Court
pKK233-2	bla, tet	Pharmacia
pAV01	pJL6 Δ(EcoRV-PvuII)::SP6 promoter	in house
pAV02	pAV01Ω(AvaI::T7 promoter)	in house
pAV1	pJL6Ω[Δ(NdeI-HindIII)::NdeI-ApaI-HindIII cloning site]	in house
pAV2	pAV02Ω[Δ(NdeI-HindIII)::NdeI-SfiI-HindIII cloning site]	in house
pAV3	pAV02Ω[Δ(NdeI-HindIII)::NdeI-AvrII(MaeI)-HindIII cloning site]	in house
pAV4	pAV02Ω[Δ(NdeI-HindIII)::NdeI-BanII-AvaI-HindIII cloning region]	in house
pAV5	pAV02Ω[Δ(NdeI-HindIII)::NdeI-AvrII(MaeI)-HindIII cloning site]	in house
pAV6	pAV02Ω[Δ(NdeI-HindIII)::NdeI-SfiI-HindIII cloning site]	in house
pAV7	pAV02Ω[Δ(NdeI-HindIII)::NdeI-StyI-HindIII cloning site]	in house
pAV8	pAV02Ω[Δ(NdeI-HindIII)::NdeI-XmaI-ApaI-HindIII cloning site]	in house
	Collagen Analogue Cassettes	
pAC91-93; pAC96-101; pAC219-223	pAV1Ω(ApaI::collagen analogue gene cassette, 9 bp repeat, 70-300 bp)	in house
pAC102-108; pAC110-116	pAV2Ω(SfiI::collagen analogue gene cassette, 9 bp repeat, 70-350 bp)	in house
pAC118-119	pAV6Ω(SfiI::collagen analogue gene cassette, pAC103)	in house
pAC120-124	pAV6Ω(SfiI::collagen analogue gene cassette, 54 bp repeat, 100-400 bp)	in house
pAC129	pAC120Ω(BamHI::cI857, 2.4 kb BglII) Δ collagen analogue gene (375—200 bp)	in house
pAC130	pAC120Ω(BamHI::cI857, 2.4 kb BglII) cI transcription directed away from cloning site	in house
pAC131	pAC120Ω(BamHI::cI857, 2.4 kb BglII) cI transcription directed toward cloning site	in house
	Glue Decapeptide Analogue Cassettes	
pAGI-pAG8	pAV7Ω(StyI::glue decapeptide analogue gene cassette, 30 bp repeat, 150-600 bp)	in house
	Elastin Analogue Cassettes	
pAE203-215	pAV1Ω(ApaI::elastin analogue gene cassette, 15 bp repeat, 140-350 bp)	in house
	Tandem and Co-cassettes	
pAC95	pAV1Ω(ApaI::2 collagen analogue gene cassettes, 9 bp repeat, 200 + 220 bp)	in house
pAC109	pAV2Ω(SfiI::2 collagen analogue gene cassettes, 9 bp repeat, 120 + 200 bp)	in house
pAC117	pAV2Ω[SfiI::multiple collagen analogue gene cassettes, 9 bp repeat, 250 + 2-3x (100-190) bp]	in house
pAC125	pAV6Ω(SfiI::4 collagen analogue gene cassettes, pAC120, 4x375 bp)	in house
pAE224, pAE226	pAV1Ω(ApaI::2 elastin analogue gene cassettes, pAE203, 2x240 bp)	in house
pACE229	pAV1Ω(ApaI::1 collagen analogue gene cassette, pAC99, 1 elastin analogue gene cassette, pAE203, 300 + 240 bp)	in house
	Expression Studies	
pAC1	pJL6Ω(ClaI::collagen analogue gene, 335 bp)	in house
pAC3	pKK233-2Ω(NcoI::collagen analogue gene, pACI, 335 bp)	in house

Table IX. Constructed Gene Cassette Expression Vectors[1]

Designation	Nde I-Hind III Cloning Region	Cloning Site	Derived From
pAV1	Met Gly Pro CATATGGGGCCCTA ACCCCGGGATTCGA	Apa I	pJL6
pAV2	Met Gly Pro Pro Gly Pro TATGGGGCCGCCAGGGCCGTA ACCCCGGCGGTCCCGGCATTCGA	Sfi I	pAV02
pAV3	Met Pro Arg TATGCCTAGGTA ACGGATCCATTCGA	Avr II/Mae I	pAV02
pAV4	Met Pro Ser Pro Pro Pro Ser TATGCCGAGCCCGCCCCCGAGCTA ACGGCTCGGGCGGGGGCTCGATTCGA	Ban II/Ava I	pAV02
pAV5	Met Ala Asn Ile Asn Asn Arg Pro Arg TATGGCTAACATTAACAACCGTCCTAGGTA ACCGATTGTAATTGTTGGCAGGATCCATTCGA	Avr II/Mae I	pAV02
pAV6	Met Ala Asn Ile Asn Asn Arg Gly Pro Pro Gly Pro TATGGCTAACATTAACAACCGTGGGCCGCCAGGGCCGTA ACCGATTGTAATTGTTGGCACCCGGCGGTCCCGGCATTCGA	Sfi I	pAV02
pAV7	Met Ala Lys Ala TATGGCCAAGGCTTA ACCGGTTCCGAATTCGA	Sty I	pAV02
pAV8	Met Gly Pro Pro Gly Pro Pro Gly Pro Pro TATGGGTCCGCCCGGGCCGCCGGGGCCCCCGTA ACCCAGGCGGGCCCGGCGGCCCCGGGGGCATTCGA	Xma I/Apa I	pAV02

1. pAV1 is a 5.1 kb rop' expression vector, the others are 3.2-3.3 kb rop expression vectors.

2. Construction of Collagen, Polyphenolic Protein and Elastin Analogue Gene Cassettes

The oligodeoxynucleotides which have been employed in the construction of various synthetic gene cassettes are shown in Figure 15. Gene cassettes with highly repeating DNA sequences were generated from these DNA duplexes by slowly annealing the complementary strands so as to favor the formation of double-stranded DNA with the given offset (e.g., 1 with 3 bp 3' overhangs) and then annealing and ligating core DNA in the presence of linker DNAs (e.g., 1 in the presence of 3). Digestion of the products with the restriction enzyme whose recognition site occurs in the linker DNA and purification of higher molecular weight synthetic DNAs led to gene cassettes ready for cloning. The average length within the distribution of gene cassette sizes could be modulated by varying the molar ratio of core oligodeoxynucleotides to linker oligodeoxynucleotides. In the case of synthetic gene cassettes made with DNAs 1 and 3 of Figure 15, the synthetic DNAs were cloned into pAV2 or pAV6. This scheme was successfully applied to all gene cassette constructs. A listing of recombinant plasmids cloned during the course of this contract and the size of the synthetic genes based upon physical mapping is given in Table VIII. Included in this listing are collagen analogue gene cassettes with oligodeoxynucleotide repeat sequences of 9 base pairs (bp) or 54 bp flanked by direct DNA repeats harboring ApaI or SfiI sites, M. edulis bioadhesive protein analogue gene cassettes with a repeat complexity of 30 bp flanked by StyI sites, and elastin analogue gene cassettes with a repeat complexity of 15 bp and ApaI flanking sites. Close examination of Table VIII will reveal that the largest single synthetic gene cassette recovered during our screening was 600 bp, despite the fact that the gene cassettes prepared in vitro ranged over 2000 bp in size prior to cloning.

3. Construction and Cloning of Tandem Cassettes and Cocassettes

Having cloned a number of synthetic gene cassettes, a major objective was then to demonstrate successful cloning and expression of multiple tandem cassettes or cocassettes (mixed cassettes) primarily using collagen and sea mussel polyphenolic protein analogue gene cassettes chosen from our gene cassette library. The purpose of this effort was to demonstrate how very

large synthetic genes capable of coding for polypeptides with MW>50,000 can be created, including those block copolymer polypeptides with mixed physical properties. Although this objective has been met in principle, the actual gene cassettes used have been those encoding analogues for collagen and elastin. For example, the plasmids pAC95 and pAC109 contain two tandem collagen analogue gene cassettes oriented correctly for expression of a higher molecular weight collagen analogue and with either ApaI or SfiI linker sequences, respectively. The plasmid pAC125 harbors a 1.5 kb collagen analogue gene composed of four tandem cassettes derived from the 375 bp pAC120 gene cassette. Similarly, pAE224 contains two tandem elastin analogue gene cassettes linked by ApaI sites. A collagen-elastin hybrid analogue cocassette gene was isolated in pACE229 with the total synthetic DNA insert over 0.5 kb in length.

SYNTHETIC GENE EXPRESSION STUDIES

The host *E. coli* used for most of the expression studies which have been completed was DC1139A (Table VIII). This strain contains a defective lambda prophage with a gene coding for a temperature-sensitive form of the cI protein which represses transcription of the p_L promoter. When the cells are grown at 30°C, transcription from p_L is almost completely repressed. The promoter becomes partially induced upon culture shift-up at 37°C and fully induced at 41°C. Many gene expression experiments were performed with the plasmid pAC1, which can produce a polypeptide which contains approximately 33 repeats of the collagen subunit tripeptide Gly-Pro-Pro. The results of some of these experiments are presented in the manuscript submitted for publication (see below) and are only briefly summarized here.

Northern blot analyses showed that large amounts of mRNA are synthesized at 41°C from pAC1 which specifically hybridize to collagen analogue sequence-specific DNA. A protein with an apparent molecular weight of 22 kDa was observed in an *in vitro* coupled transcription-translation experiment which was specifically expressed from pAC1 and which was sensitive to collagenase. *In vivo* labeling experiments using [^{14}C]proline were carried out with DC1139A as host. After preincubation at 41°C for 20 minutes, a protein band of 22 kDa was visible only with pAC1-containing cells. This protein was absent when the preincubation was carried out at 30°C as well as for pJL6-containing cells.

As a comparative study, expression of collagen analogue peptide was also evaluated using the trc promoter. For these experiments, the NdeI-HindIII insert DNA fragment from pAC1 was cloned into the NcoI cloning site of the expression plasmid pKK233-2 (Pharmacia) using NcoI linkers, giving rise to the expression plasmid pAC3. Expression from the trc promoter was induced by incubating cultures in the presence of 1 mM IPTG for at least 15 minutes. A protein of about 22 kDa was again detected uniquely from pAC3 in the presence of IPTG. The level of expression in the p_L and trc promoter systems was approximately equal.

Based on the size of the insert in pAC1 (335 bp), the molecular weight of the expected gene product should have been about 12 kDa instead of 22 kDa. We wished to confirm the identity of the 22 kDa protein as the product of the synthetic collagen analogue gene. A Western blot was performed using rabbit antiserum raised against chemically synthesized $(Pro\text{-}Pro\text{-}Gly)_{14}$ conjugated to bovine serum albumin. The immunizations were carried out and the immune antisera were collected over a three month period at Berkeley Antibody Co., (Richmond, CA). An indirect immunoassay method was used to detect (Pro-Pro-Gly)-specific proteins with the secondary antibody being goat anti-rabbit IgG coupled to horseradish peroxidase. The data are shown in Figure 16. A major band at approximately 22 kDa was seen only in the lanes containing cellular extracts where pAC1 had been present. This confirms that the 22 kDa protein observed in in vitro and in vivo labeling experiments with pAC1 is the collagen analogue polypeptide. Furthermore, the biologically produced collagen analogue and the chemically synthesized poly(Pro-Pro-Gly) must share antigenic domains, indicating that collagen analogue polypeptides from these two sources have a similar structural conformation.

Experiments were also performed to investigate the intracellular stability of the collagen-like polypeptide synthesized from pAC1. Cell cultures were pulse-labeled with [^{14}C]proline and chased with an excess of unlabeled proline for extended periods of time. Analysis of the data showed that about 50% of the labeled collagen polypeptide was degraded within 20 minutes in the E. coli host strain DC1139A. When the experiment was repeated in an engineered protease-deficient mutant (htpR$^-$) designated IG110, no detectable loss of collagen analogue polypeptide occurred over a two hour period. The heterologous protein encoded by pAC1 is therefore most likely susceptible to degradation by the Lon protease or another intracellular protease controlled

by the *hptR* gene. As a control, stability of the collagen analogue polypeptide was also assayed in the *htpR*$^+$ parent strain IG109 from which IG110 was derived. The data were similar to those obtained for DC1139A. The results illustrate the importance of protease-minus strains in maximizing the yield of heterologous collagen analogue polypeptides and possibly other polypeptides obtained from internally repetitive synthetic gene cassettes such as those encoding analogues to the *M. edulis* polyphenolic protein.

STABILITY OF SYNTHETIC GENE CASSETTES

1. Gene Cassette Composition

One of our major biological objectives during the current contract period was to assess the stability of gene cassettes composed of repeating DNA sequences. Collagen analogue gene cassettes containing 9 bp repeat units manifest a significant level of insert instability when either large scale plasmid preparations or plasmid mini-preparations were isolated from growing recombinant cultures. This synthetic gene instability could be demonstrated either as differences in insert size for replicate isolates from parallel cultures (Figure 17) or as multiple insert DNA bands during physical mapping of single isolates (Figure 18). Increasing size of synthetic gene cassettes was directly correlated to gene cassette instability. For example, it is noteworthy that size analysis of *in vitro* constructed gene cassettes repeatedly showed lengths ranging up to several kb, but the largest characterized single cloned collagen analogue gene cassettes established in *E. coli* and stored in our recombinant clone collections are no larger than 350-400 bp (see Table VIII). A few collagen analogue gene cassettes have been observed during restriction enzyme mapping which were 0.5-0.6 kb in length, but these inserts were either lost or a population of smaller cassettes accumulated during further growth. These phenomena were particularly evident for gene cassettes greater than 300 bp in length, although accumulation of partially deleted cassettes was also observed for smaller collagen analogue gene cassettes (e.g., Figure 23, lane 9).

Repetitive gene instability undermines our ability to successfully establish genetically engineered recombinant bacterial strains expressing high levels of high molecular weight *M. edulis* polyphenolic protein analogue or

collagen analogue polypeptides. We were hopeful early in our efforts that such events would occur less frequently as the diversity (genetic complexity) of the repeating genes was increased. We designed a 54 base oligonucleotide collagen analogue repeat from which a more complex collagen analogue gene cassette could be constructed using overlapping synthetic DNAs (2 in Figure 15). These DNAs employ alternative favored codons for Gly and Pro and encode six unique (Gly-Pro-Pro) repeats. These oligonucleotides were used to construct and clone SfiI gene cassettes. Plasmids carrying such diversified collagen analogue gene cassettes still showed deletion events upon characterization, this time at 54 bp intervals with less frequent events which were separated by 18 or 36 bp (Figure 19). These latter deletions correspond to several different nested repeats that are present in the diversified sequence. Despite the instability of these gene cassettes, multiple cassettes could still be linked in tandem to generate very long repeating synthetic collagen analogue genes (Figure 20). The larger gene pictured in Figure 20 was very unstable, with most of the collagen analogue genes from the plasmid pool being less than the full 1.5 kb in length. Culturing of single collagen analogue cassettes can lead to decreases in the size distribution of deleted cassettes down to total or near total deletion (Figure 21). As a result of our observations with synthetic collagen analogue gene cassettes, we have concluded that preferential codon usage is a minor consideration relative to maximizing sequence diversity when designing stable forms of repeating synthetic gene sequences.

A legitimate concern is whether the instability of collagen analogue gene cassettes is generally representative for all repetitive analogue gene cassettes. This concern is addressed in Figure 22 wherein the accumulation of partially deleted cassettes is assayed from large-scale cultures of strains carrying single analogue gene cassettes of collagen or M. edulis polyphenolic protein, tandem analogue cassettes of elastin, and a collagen-elastin cocassette. Since all of these analogue gene cassettes show some instability, the lack of collagen analogue gene cassette stability can not be merely ascribed to peculiarities in its base composition or sequence. However, significant differences in analogue cassette stability were apparent with elastin, M. edulis polyphenolic protein, and collagen showing, respectively, low, intermediate, and high accumulation of partially deleted cassettes. The tandem addition of a collagen to an elastin cassette also increased the level

of accumulation of shorter inserts compared to that observed for two tandemly cloned elastin cassettes. Thus, we were actually fortunate to have chosen the relatively unstable collagen analogue as a model system since we could detect gene deletions so readily. The overall stability observed for each of these constructs is a composite of several variables including base composition, base sequence, repeat size, cassette/insert size, and parent replicon. Further work is required to sort out the individual contribution of each variable to overall cassette stability.

2. Basal Transcription and Synthetic Gene Cassette Instability

Several models might explain our results on synthetic gene cassette instability. One of these models links transcriptional activity to gene deletions. The ColE1-derived expression plasmids with which we have worked are known to replicate at a high level in E. coli and are present in relatively high copy number per cell (20-30 replicons/cell). Our development of rop$^-$ expression plasmids further increases the plasmid copy number per cell. It is known that a single cI gene leads to the accumulation of only 100-200 cI repressor molecules per cell, which means that in our best-studied regulated gene expression host-vector combinations there is as little as 5 or less cI repressor molecules per recombinant replicon. Induction of lambda prophage using the same thermosensitive cI857 repressor gene which we have used has shown that induction begins when the number of active cI857 repressor molecules drops below about 10 repressor molecules per p_L promoter site. Therefore, some basal level of transcription is to be expected in our host-vector system from the p_L promoter controlling transcription of our synthetic gene cassettes. It is possible that such adventitious transcription of any of the synthetic gene cassettes which we have constructed would contribute in some fashion to the gene deletion events. We tested this hypothesis by transferring a 2.4 kb BglII fragment of lambda DNA containing an entire cI857 gene with its promoter elements into the unique BamHI site of pAC120 to generate pAC130 and pAC131 (Table VIII). Some isolates were obtained during the construction of pAC120 derivatives which are partially deleted in the synthetic gene cassette; one of these (pAC129) was tested to see if there was any connection between gene cassette size, basal transcriptional activity and gene deletions. Large scale cultures of E. coli

strain DC1138 containing each plasmid were grown at 30°C, 37°C or 41°C in order to compare levels of gene deletions. Negative results were obtained from these experiments since overrepression of transcription from the p_L promoter did not inhibit the process of gene cassette deletions (Figure 23).

3. Host Effects on Gene Cassette Stability

Two different types of genetic background for the E. coli hosts have been examined for influence on gene cassette stability. One genetic background was defective in the RecA gene product which is required for homologous recombination by the RecBCD and RecF pathways [Smith, G.R., Microbiological Reviews, 52, 1-28 (1988)]. We tested a number of different recA mutant alleles and found that none showed a stabilizing effect on gene cassettes containing repetitive DNA. It should be noted that strain constructs made during the course of this contract carry the stringent recA mutation (srlR-recA)306::Tn10 which is a deletion of all but about 50 bp of the recA gene.

The other host genetic background that was examined was the lambda red-mediated homologous recombination system. The E. coli host strain DC1138 used for cloning synthetic gene cassettes carries a functional but repressed red system; the red gene system in DC1138 is also presumably uninducible due to the recA mutation. The red genes of the prophage carried in gene expression strains DC1139A and IG110 are deleted. Figure 18 shows that in either the red^- or red^+ host substantial levels of deleted collagen analogue gene cassettes accumulated upon culturing. Thus, the lambda red recombination system is not the primary mediating factor of bacterial hosts which contributes to the instability of repetitive DNA gene cassettes.

POSSIBLE MECHANISMS OF SYNTHETIC GENE CASSETTE INSTABILITY

It is clear that the gene cassettes containing repetitive DNA which we have constructed are unstable in E. coli. Our observations prompt consideration of possible mechanisms for recombination which determines instability in our system. Several possible mechanisms have not been eliminated by our investigations. One of the most obvious is the slipped mispairing model of single-stranded DNA [G. Streisinger, C.S.H.S.Q.B. 31,

77-84 (1966)]. A major obstacle of applying this model to our data is that it predicts an equal likelihood of product plasmids containing insertions or deletions of repetitive DNA. Although some low level of enlarged gene cassettes has been observed in some experiments, it is clear that synthetic gene deletions are overwhelmingly favored in our system. On the other hand, the Streisinger model can be defended by arguing that since the original plasmid and its deletion products belong to the same incompatibility group, any selective advantage for replication or maintenance of deleted product plasmids provides a means for a growing bacterial population to accumulate deletions within the plasmid pool [cf. R.P. Novick, Microbiological Reviews 51, 381-395 (1987)]. The tendency for accumulation of plasmids with smaller synthetic gene cassettes and of complete deletion of gene cassettes upon continued culturing of recombinant E. coli populations provides a hint that such an advantage does exist. If slipped mispairing underlies the deletion phenomenon, then one would predict that by increasing synthetic gene cassette DNA complexity (and thereby the distance between homologous regions on complementary strands), one could reduce the rate and extent of accumulation of deleted gene cassettes within the plasmid pool during culturing.

Slipped mispairing is not the only model which can account for repetitive gene cassette instability. Single-stranded gene cassette DNA generated during replication or repair processes, or formed as the result of topological or torsional stress due to plasmid supercoiling, could serve as recombinogenic substrates for any of a number of gene products of the RecBCD, RecE and/or RecF recombination pathways. However, the generation of single-stranded DNA may not necessarily be a prerequisite for gene cassette deletions. Based on model studies of recombination of completely intact homologous duplexes, deletions can occur by a pair-first pathway [cf. J.H. Wilson, P.N.A.S. 76, 3641-3645 (1979)]. The importance of this plausible theoretical pathway of recombination for in vivo deletions of synthetic gene cassettes containing repetitive DNA is uncertain. Such a pathway is presumably independent of recA gene function [e.g., see G.R. Smith, Microbiological Reviews 52, 1-28 (1988) for a review of biochemical reactions of RecA protein]. However, plasmid recombination by the RecE pathway is also independent of recA or recF gene functions (R.A. Fishel et al., Nature 294, 184-186 (1981)]. Intraplasmid recombination is reduced as much as 100-fold in most recA$^-$ strains [A. Laban and A. Cohen, Mol. Gen. Genet. 184, 200-207 (1981)]. One possibility is,

therefore, that part of any residual intraplasmid recombination activity in recA⁻ hosts is due to basal level expression of the repressed RecE pathway. Although this is unlikely, since no difference in the level of intraplasmid recombination has been observed among recA⁻ strains which contain a repressed RecE recombinational pathway compared to those deleted (rac⁻) for the RecE pathway (A. Laban and A. Cohen, op. cit.), only continued experimentation will determine whether or not the RecE pathway or other above-mentioned models are operational in our system.

PUBLICATIONS SUBMITTED DURING THE COURSE OF THE CONTRACT

1. Swerdloff, M.D., Anderson, S.B., Sedgwick, R.D., Gabriel, M.K., Brambilla, R.J., Hindenlang, D.M., and Williams, J.I., "Solid Phase Synthesis of Bioadhesive Analogue Peptides with Trifluoromethanesulfonic Acid Cleavage from PAM Resin" (submitted to Int. J. Peptide Prot. Res.).

2. Goldberg, I., Salerno, A.J., and Williams, J.I., "Cloning and Expression of a Synthetic Collagen Analogue Gene in Escherichia coli (submitted to Gene).

POSTERS PRESENTED AT CONFERENCES

1. Brambilla, R., and Swerdloff, M., "High Resolution 100 MMz Carbon-13 NMR Spectral Studies of Peptides Attached to Solvent Swollen Merrifield Resins" (28th Experimental NMR Conference, June, 1987).

2. Swerdloff, M., and Brambilla, R., "Side Reactions During the Solid Phase Synthesis of Trifluoromethanesulfonic Acid Cleavage of Peptides from PAM Resins" (Tenth American Peptide Symposium, August, 1987).

3. Gabriel, M.K., Hindenlang, D.M., Sedgwick, R.D., Swerdloff, M.D., and Williams, J.I., "Characterization of Synthetic Adhesive Decapeptides by FAB-MS and HPLC-MS (36th ASMS Conference on Mass Spectrometry and Allied Topics, June, 1987).

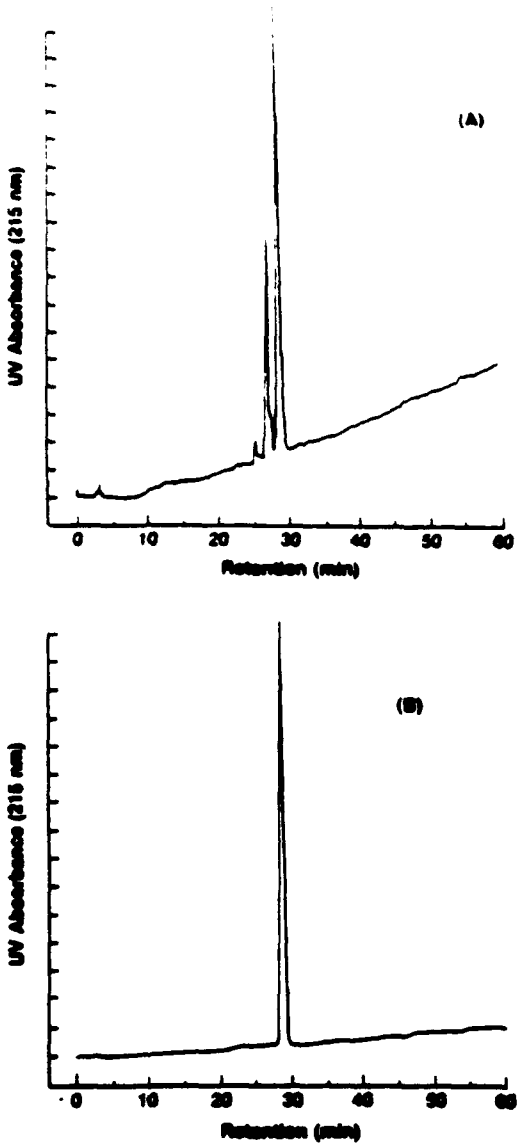


Figure 1. HPLC chromatograms of GLUE-12 decapeptide (A) crude material after the two-step TFMSA deprotection and cleavage procedure and (B) after neutralization and purification by preparative HPLC as described in the text.

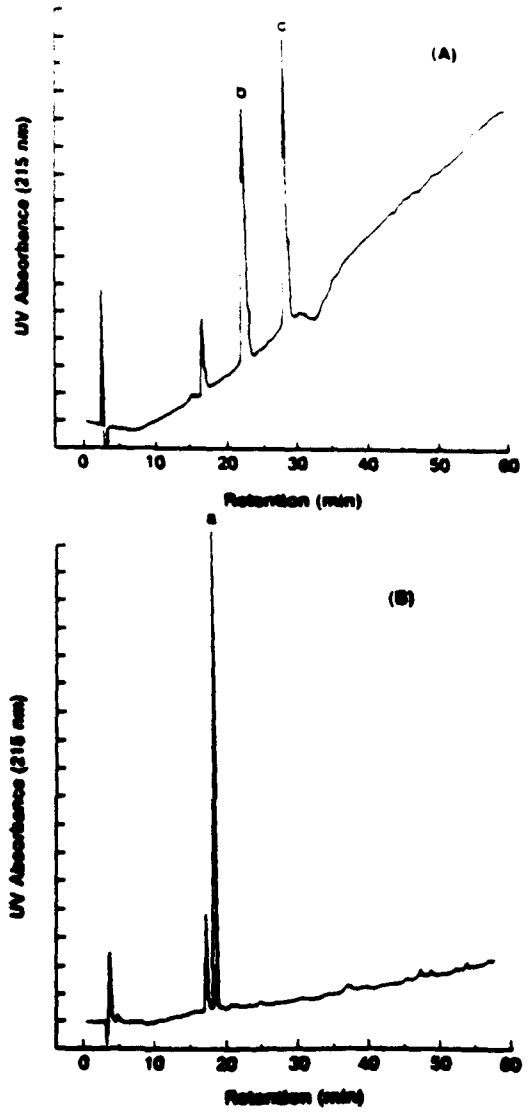


Figure 2. HPLC chromatograms showing the progression of deblocking the ϵ-amino groups of purified GLUE-12. The chromatogram in panel A is of GLUE-12 after 18 hr incubation with the deblocking solution (piperidine:methanol:water, 20:70:10) and shows a mixture of partially deprotected material (peak b) and fully protected GLUE-12 (peak c). Panel B shows fully deprotected material (peak a) after 96 hr incubation with the deblocking solution. The fully deblocked GLUE-12 was then desalted and used in oxidation experiments.

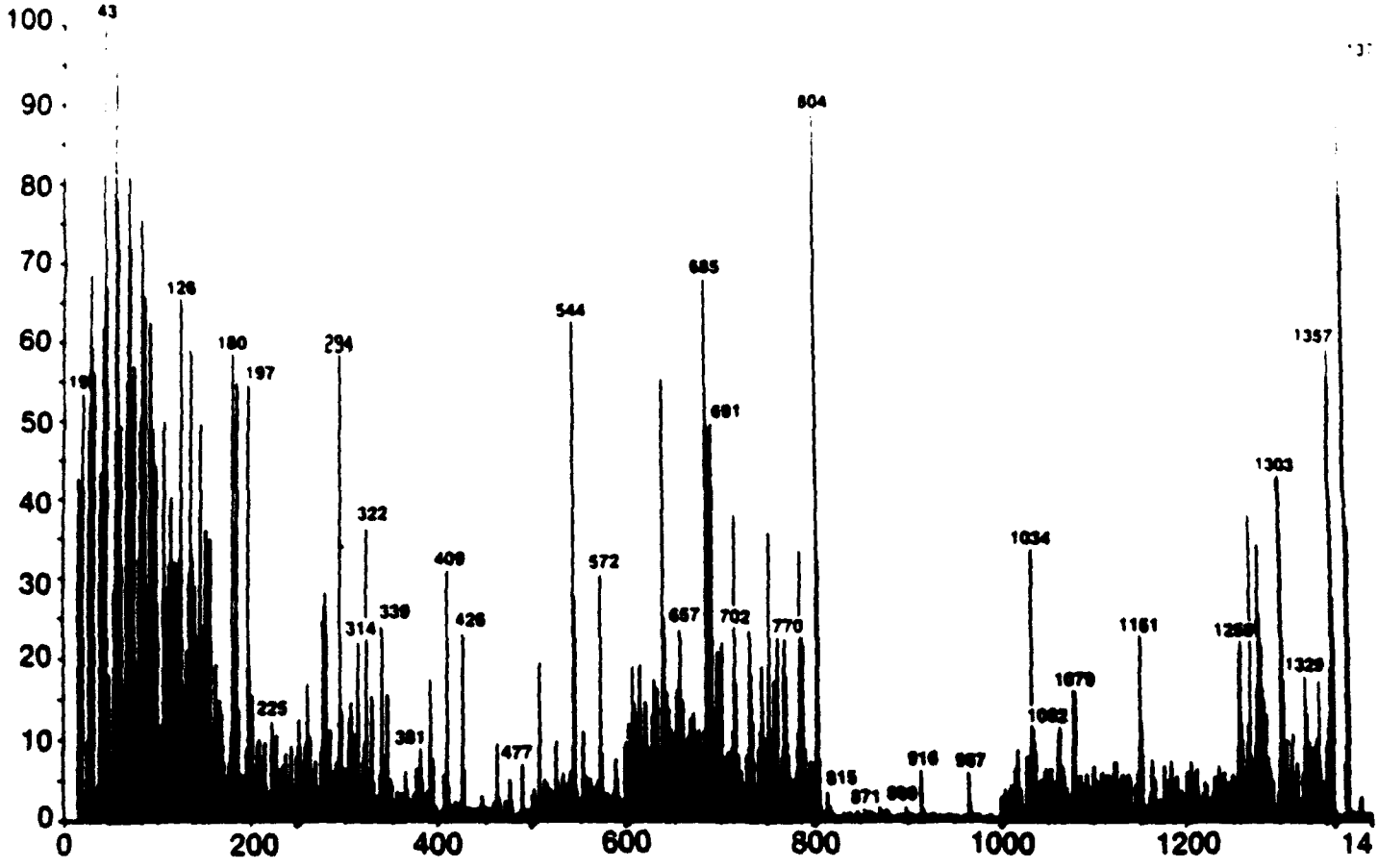


Figure 3. FAB-MS of purified GLUE-12. Scale expansions were as follows: 0-250, X1; 250-500, X2; 500-600, X10; 600-800, X40; 800-1000, X4; 1000-1360, X50; 1360-1400, X4. The M+ line at 1376 is exact for GLUE-12. The complete sequence can be constructed by matching the mass lines in this figure with the 'A' fragmentation series listed in Table II, and confirmed by the overlapping regions of the 'B' and 'C' series. Additional confirmation was provided by cleavages retaining the carboxy terminus which occurred as the 'Y' series.

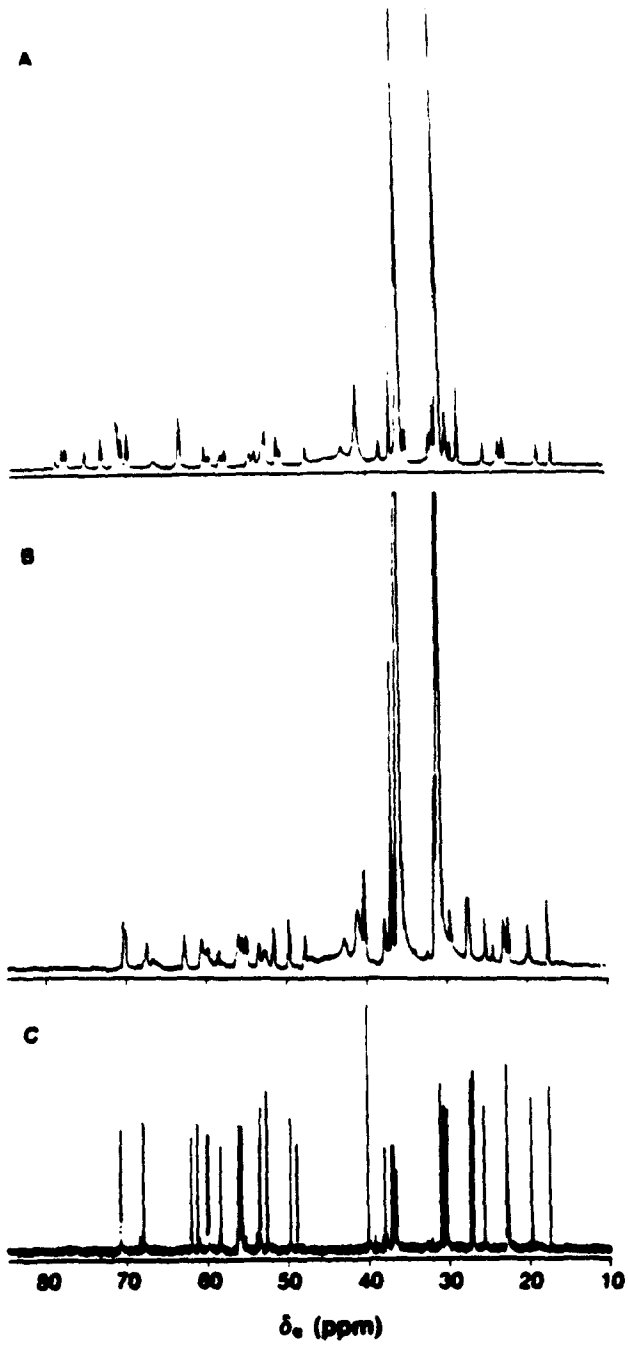


Figure 4. The high-field region (10-75 ppm) cf the ^{13}C NMR spectrum for (A) protected GLUE-2-PAM resin, (B) deprotected GLUE-2-PAM resin, and (C) deprotected GLUE-2 in D_2O.

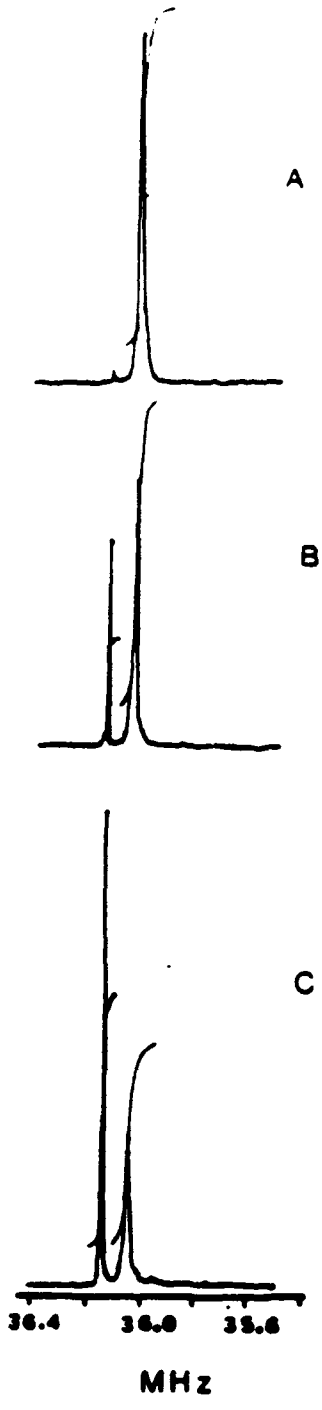
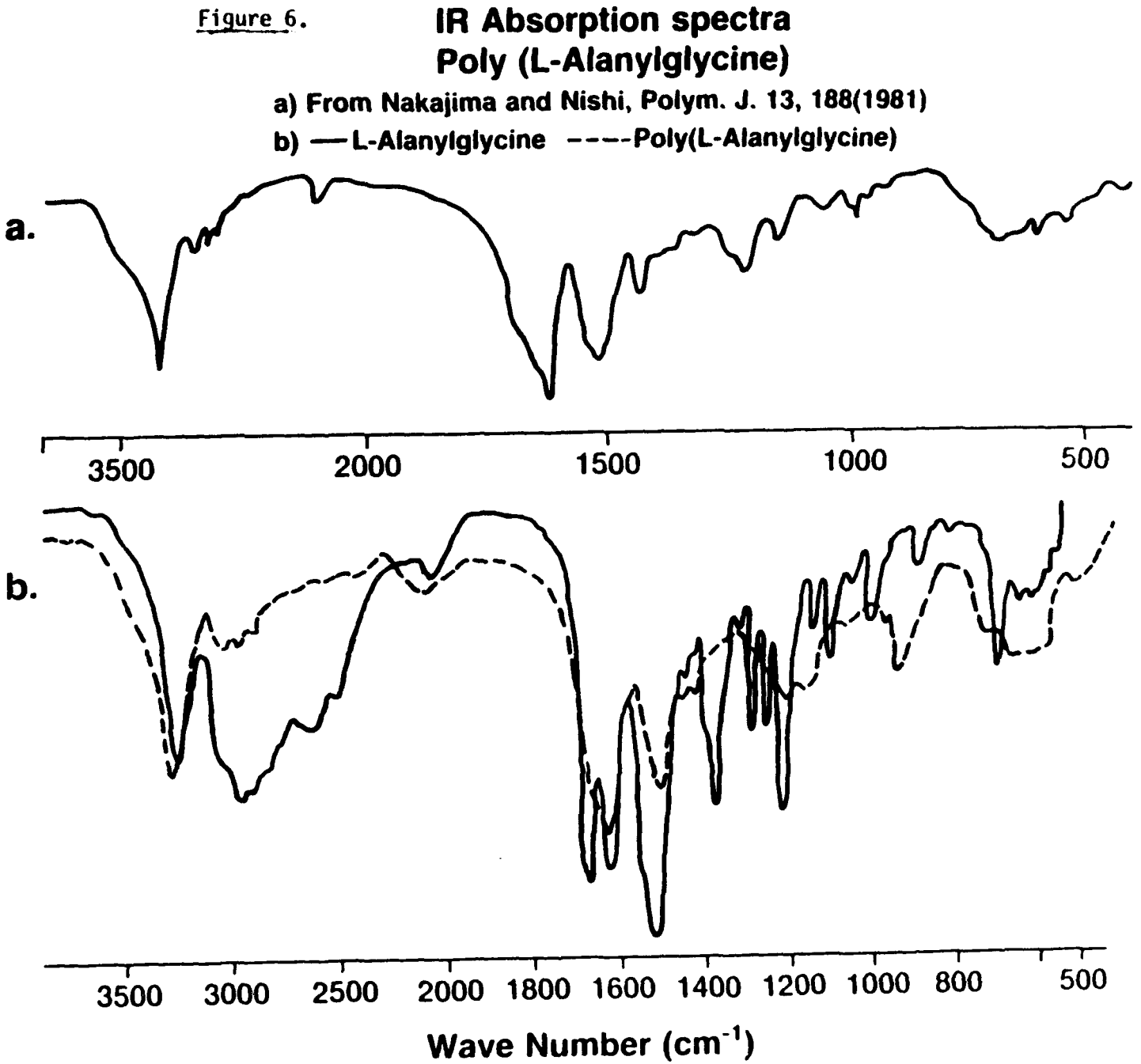


Figure 5. ^{19}F NMR study of cleavage of TFA-blocked lysines in poly(GLUE-12) with methanolic solution of piperidine. These representative NMRs show the extent of cleavage with time; (A) t= 0 min (B) t= 15 min and (C) t= 975 min. The NMR peak at 36.04 MHz (27.75 ppm) representing TFA-amides in the reaction:

$$\text{lysine-}\epsilon\text{-NH-}\overset{\overset{O}{\|}}{C}\text{-}CF_3 \xrightarrow{\text{base}} CF_3COOH + \text{lysine-}NH_2$$

was found to decrease and the peak for CF_3COO^- at 36.16 MHz (27.65 ppm) increased over time.

Figure 6.


IR Absorption spectra
Poly (L-Alanylglycine)
a) From Nakajima and Nishi, Polym. J. 13, 188(1981)
b) —L-Alanylglycine ----Poly(L-Alanylglycine)
a.
3500
2000
1500
1000
500
b.
3500
3000
2500
2000
1800
1600
1400
1200
1000
800
500
Wave Number (cm-1)

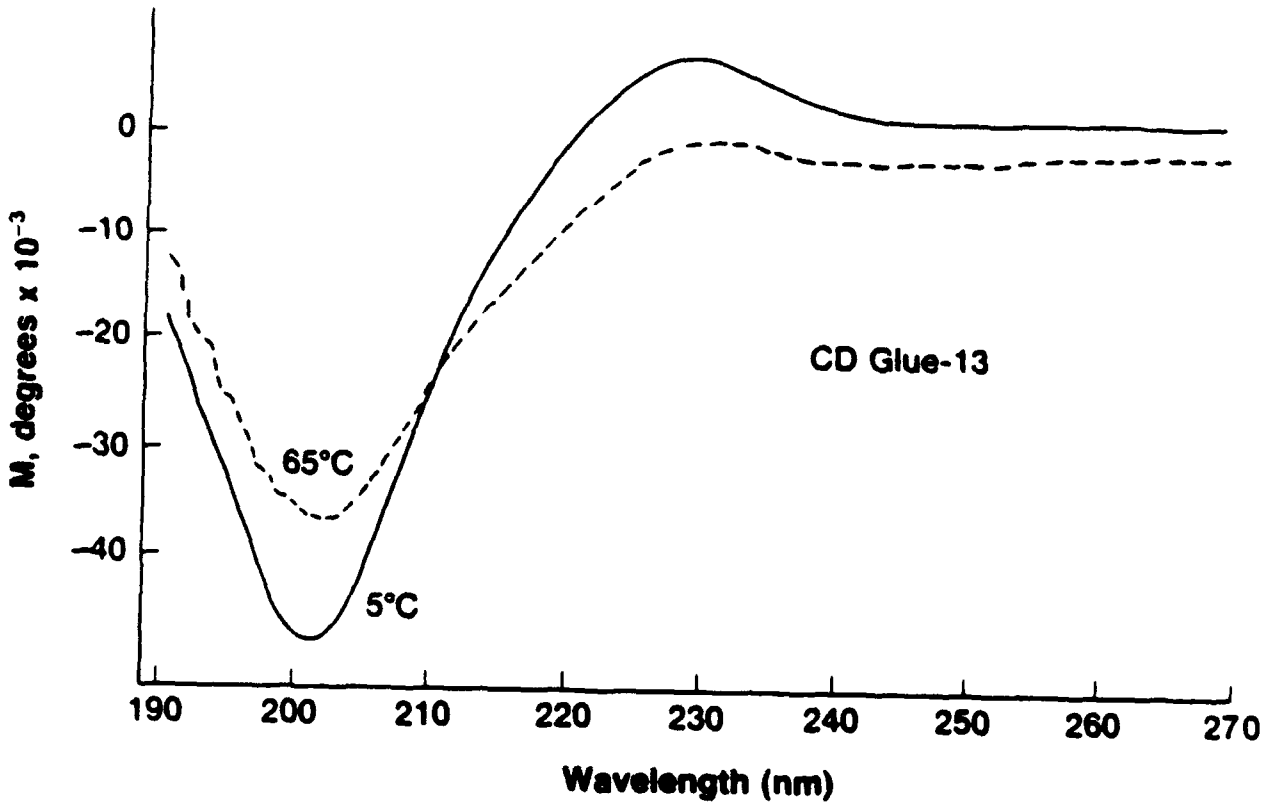
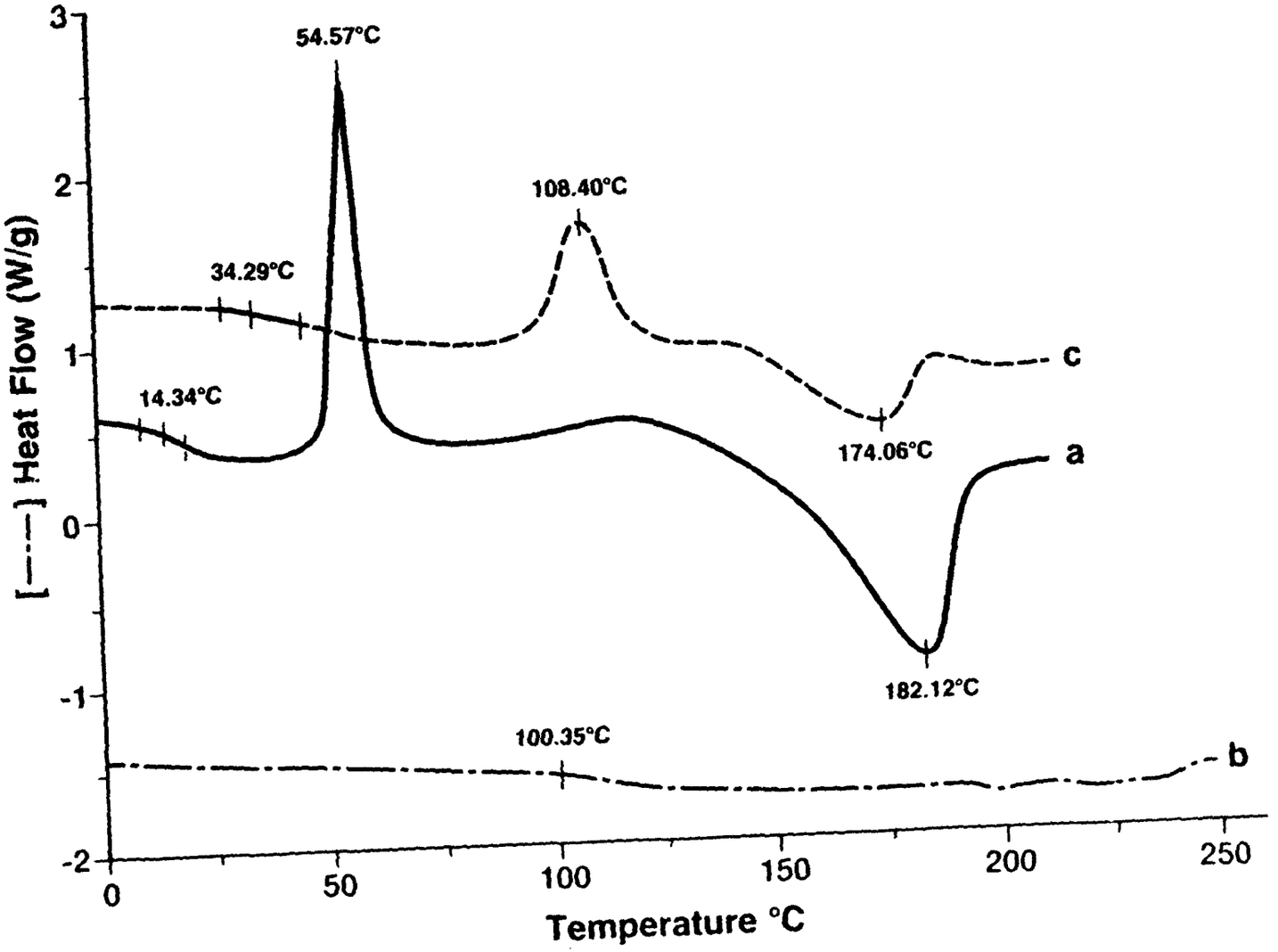


Figure 7. Circular Dichroism spectra of GLUE-13 (2 mM in water) taken at 5 ° C (———) and 65 ° C (– – –).

Figure 8. **DSC Profiles Of: a) Nylon-6 b) Polydecapeptide (PDP) and c) Nylon-PDP Block Copolymers**



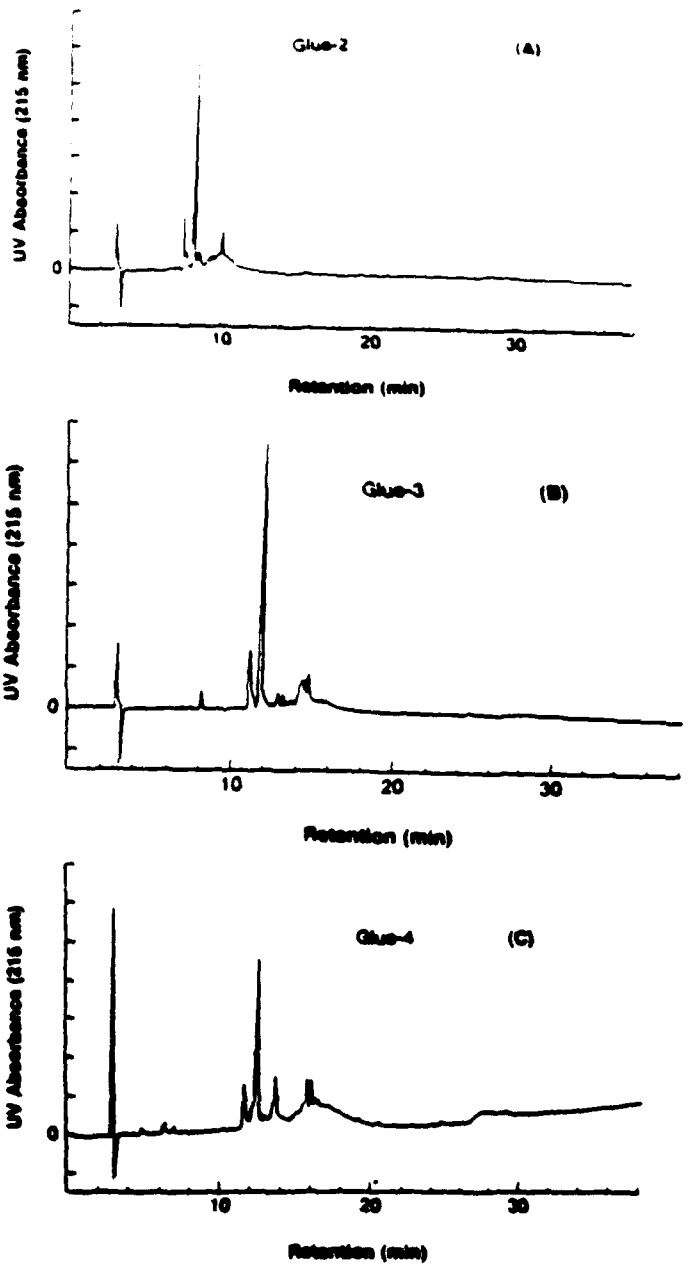


Figure 9. HPLC chromatograms of oxidation reaction mixtures of GLUE-2 (A), GLUE-3 (B), and GLUE-4 (C) after 100 minutes of incubation. 0.62 mM of each GLUE decapeptide was incubated with 50 μg (GLUE-2 and -3) or 73 μg (GLUE-4) of tyrosinase in 0.1 M phosphate buffer, pH 8.

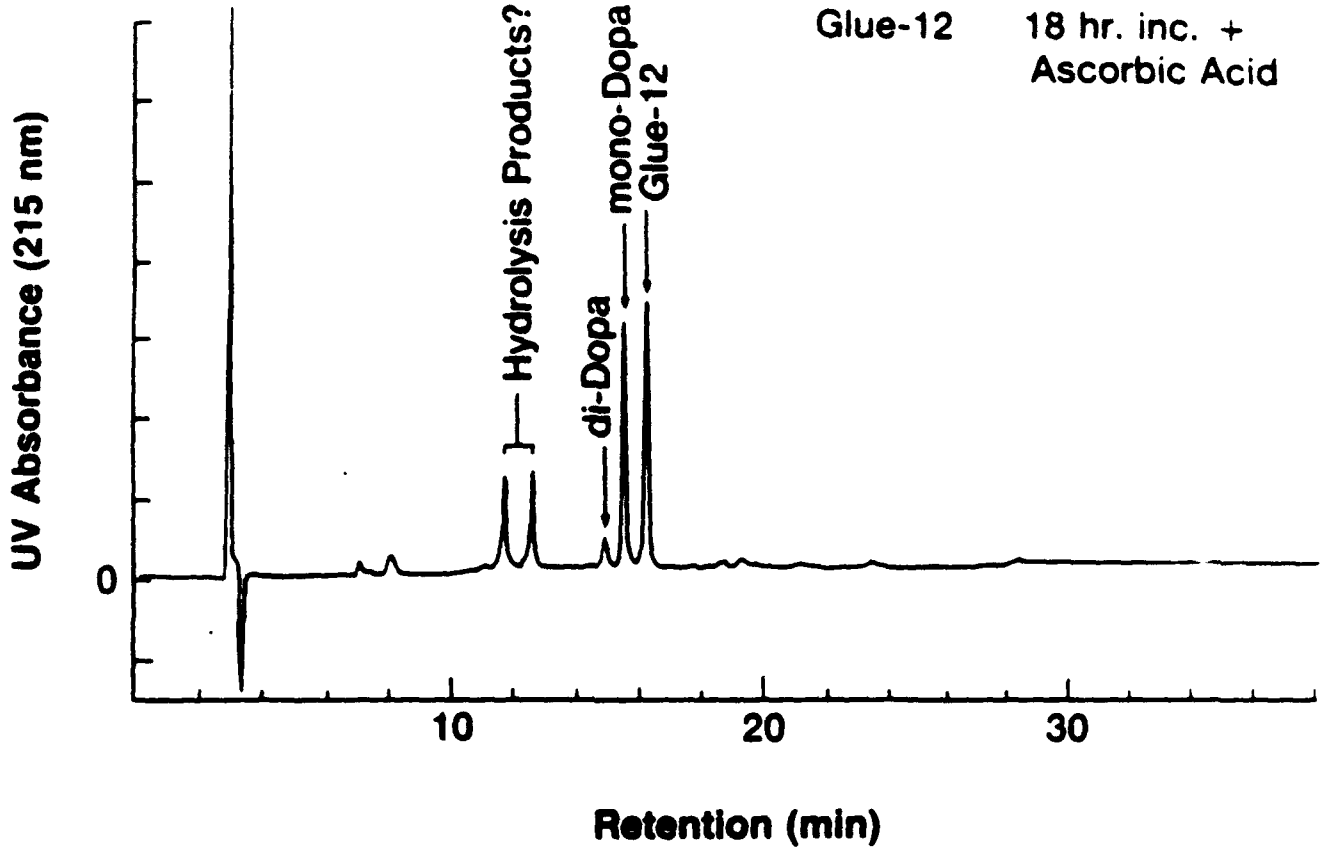


Figure 10. HPLC chromatogram of the oxidation of GLUE-12 in the presence of ascorbic acid. GLUE-12 (0.5 mM) was incubated in the presence of ascorbic acid (5 mM) and 50 μg of tyrosinase for 18 hours. Hydrolysis products (the peaks eluting between 12 and 14 minutes) were most prominent in reaction mixtures incubated for long periods of time. The mono- and di-Dopa oxidation products were identified by coupled HPLC-Thermospray mass spectrometry.

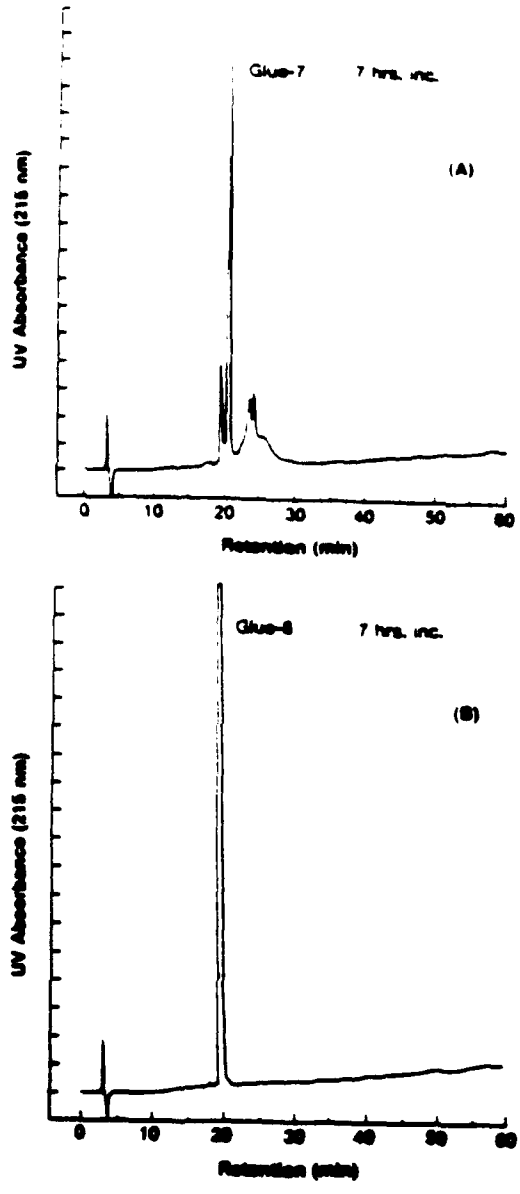


Figure 11. HPLC chromatograms from the oxidations of GLUE-7 (A) and GLUE-8 (B). GLUE-7 or GLUE-8 (0.44 mM) was incubated with 150 µg of mushroom tyrosinase for 7 hours in the presence of 0.1 M phosphate buffer, pH 8.

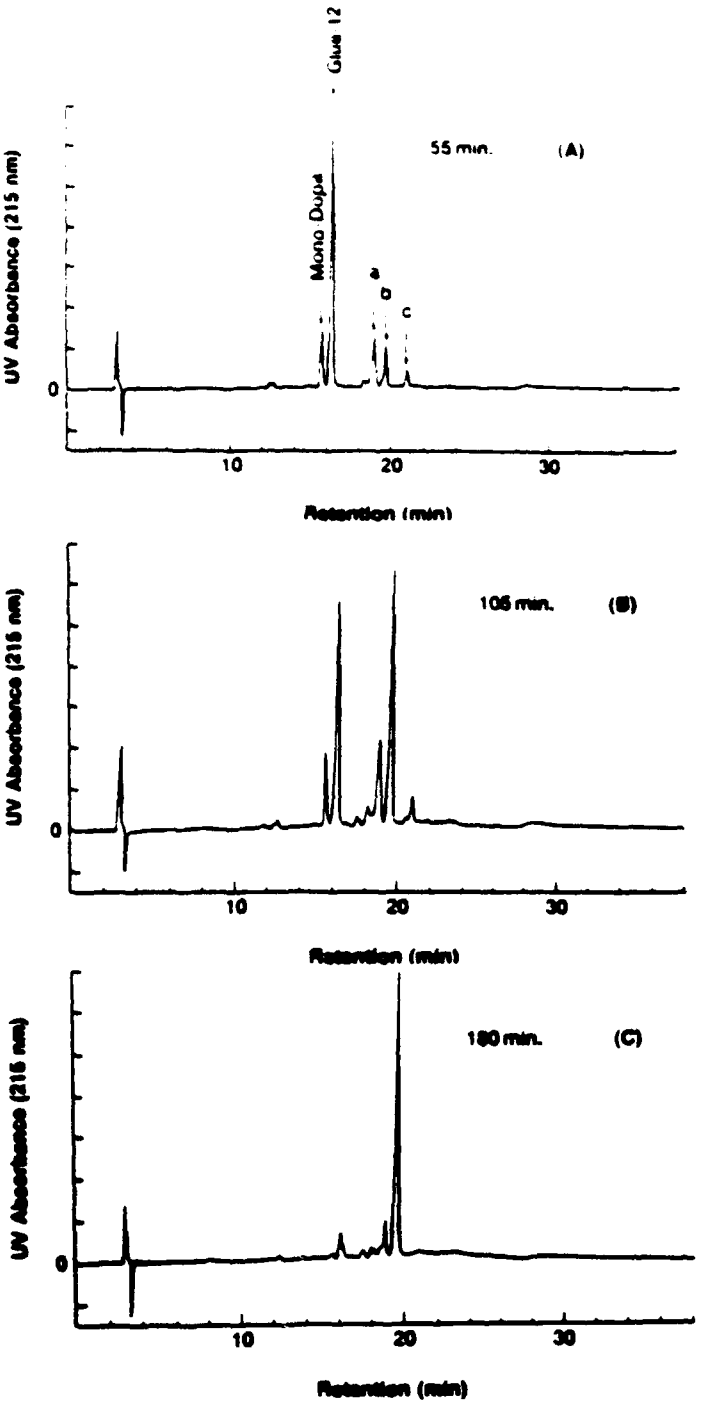


Figure 12. HPLC chromatograms from the oxidation of 0.6 mM GLUE-12 with mushroom tyrosinase (50 µg) for 55 min (A), 105 min (B), and 180 min (C) in 0.1 M phosphate buffer, pH 8. The major component in Panel A is the GLUE-12 parent compound (16.3 min retention time), which is immediately preceeded by the mono-Dopa oxidation product. Thermospray-MS analysis of the major component in the 180 min incubation mixture (19.6 min retention time, peak b in Panel A) indicated that its molecular weight was consistent with an internally crosslinked mono-quinone product. Thermospray-MS results for all the labeled peaks in panel A are listed in Table VII.

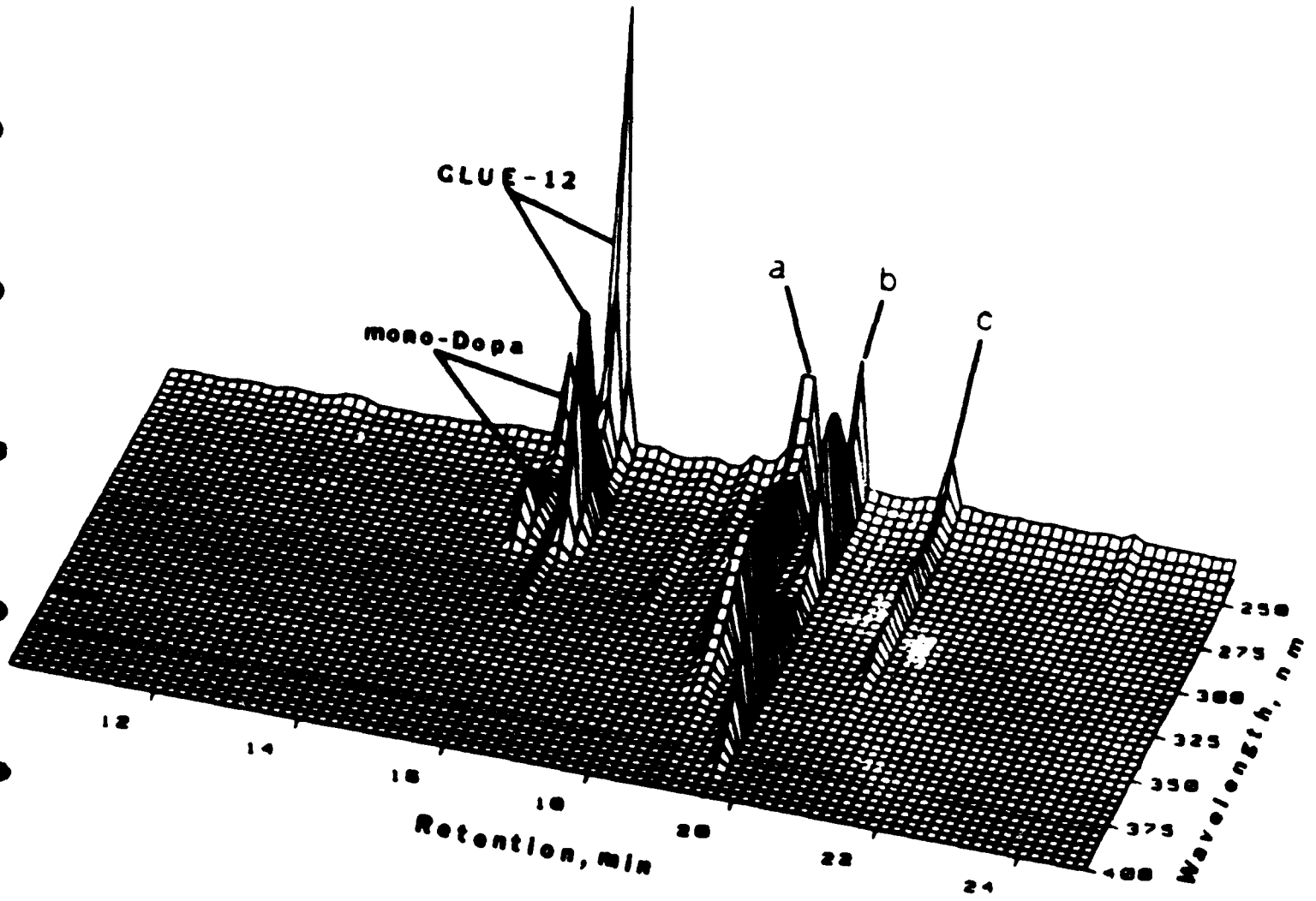


Figure 13. Results of the coupled HPLC-diode array UV/vis analysis of GLUE-12 and its oxidation products. Components eluting from the column between 10 and 25 min were scanned over the wavelength range of 230-400 nm, with the absorbance values presented as topographical elevations. Shown are the UV/vis absorbance characteristics of the GLUE-12 substrate, mono-Dopa derivative and oxidation products a, b and c. The peaks in this figure are labeled the same as in Panel A of Figure 12. As can be seen in this figure, the UV/vis absorbance characteristics of components a and c are similar to each other (λmax 280 nm, shoulder 300 nm), but distinctly different from component b (λmax 270 nm and 348 nm).

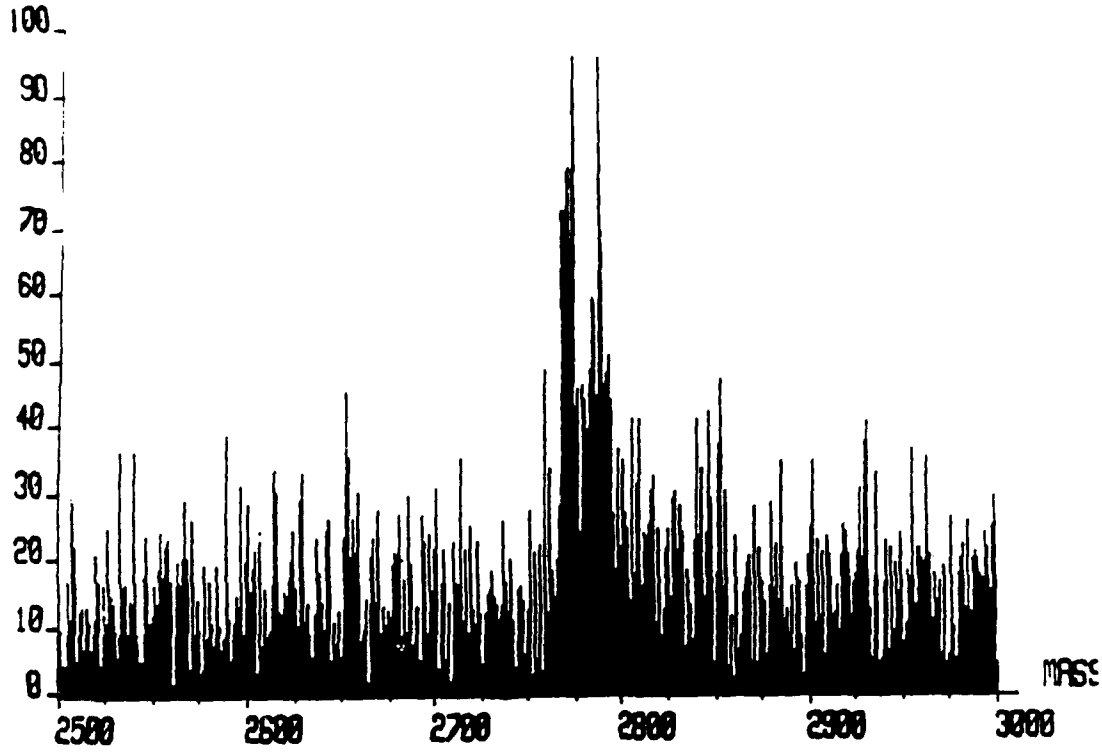


Figure 14. Coupled HPLC-Thermospray MS analysis of the poorly resolved peak of the oxidation product of GLUE-12 after long incubation. The data represent a statistical average of spectra taken at several points near the peak of the broad component. The major spectrum lines roughly approximate a dimeric form of GLUE-12. The instrument was not calibrated over this high mass range immediately preceeding the run, so there is some uncertainty in the mass assignments.

...Gly Pro Pro...

5′-GGTCCGCCGGGTCCGCCG-3′

3′-GGCCCAGGCGGCCCAGGC-5′

1 collagen, internal 9 bp repeat

...Gly Pro Pro ...

5′-GGCCCACCGGGTCCGCCAGGCCCGCCGGGTCCACCG

3′-GGCCCGGGTGGCCCAGGCGGTCCGGGCGGCCCAGGTGGC

GGCCCGCCAGGTCCGCCG-3′

CCGGGCGGTCCAGGC-5′

2 collagen, internal 54 bp repeat

Gly Pro Pro

5′-GGGCCCCCG-3′

3′-GGCCCCGGG-5′

3 collagen, elastin, Apa I linker

Gly Pro Pro...

5′-GGGCCGCCAGGGCCGCCG-3′

3′-GGCCCCGGCGGTCCCGGC-5′

4 collagen, Sfi I linker

... Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro ...

5′—CCAACCTACAAAGCTAAGCCGTCTTATCCG-3′

3′—TTTCGATTCGGCAGAATAGGCGGTTGGATG—5′

5 glue, internal 30 bp repeat

Pro Thr Tyr Lys Ala Lys Ala Ser Tyr Pro

5′—CCAACCTACAAAGCCAAGGCTTCTTATCCG—3′

3′-TTTCGGTTCCGAAGAATAGGCGGTTGGATG-5′

6 glue, Styl linker

...Gly Val Gly Val Pro...

5′-GGTGTTGGTGTTCCG-3′

3′—GGCCCACAACCACAA—5′

7 elastin, internal 15 bp repeat

Fig. 15. Oligonucleotides comprising analogue gene cassettes.

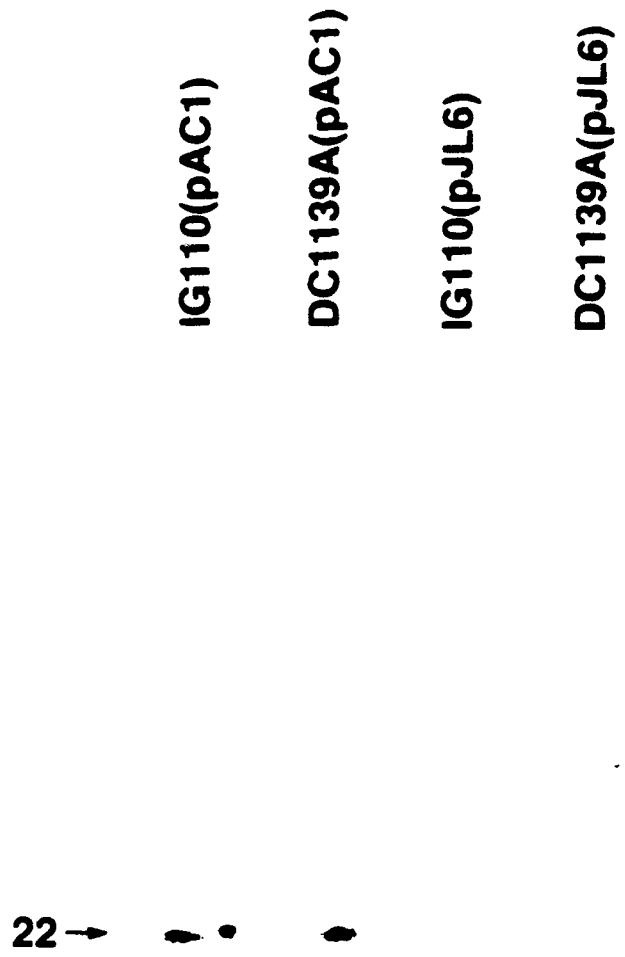


Figure 16. Immunoblotting of collagen analogue peptide produced in E. coli strains IG110 and DC1139A. The Western blot procedure was adapted from one supplied by Biorad. Cellular proteins were electrophoresed on a 12.5% SDS polyacrylamide gel and transferred to nitrocellulose in a trans-blot cell (Biorad). After transfer, the nitrocellulose was incubated in Tris-buffered saline (TBS) with 1% bovine serum albumin to reduce nonspecific antibody binding. The anti-(Pro-Pro-Gly) antiserum (1:25 dilution) was added in blocking buffer (5% nonfat dry milk in TBS) overnight. The nitrocellulose was then washed in TBS three times for 30 min. The secondary antibody (goat anti-rabbit IgG coupled to horseradish peroxidase) was added in blocking buffer and incubation was continued for 2 hours. After washing the nitrocellulose in TBS three times for 30 min., the blot was developed using diaminobenzidine as enzyme substrate. The location of the collagen analogue peptide is indicated by the arrow. All other bands are due to nonspecific activity of the antiserum against E. coli proteins.

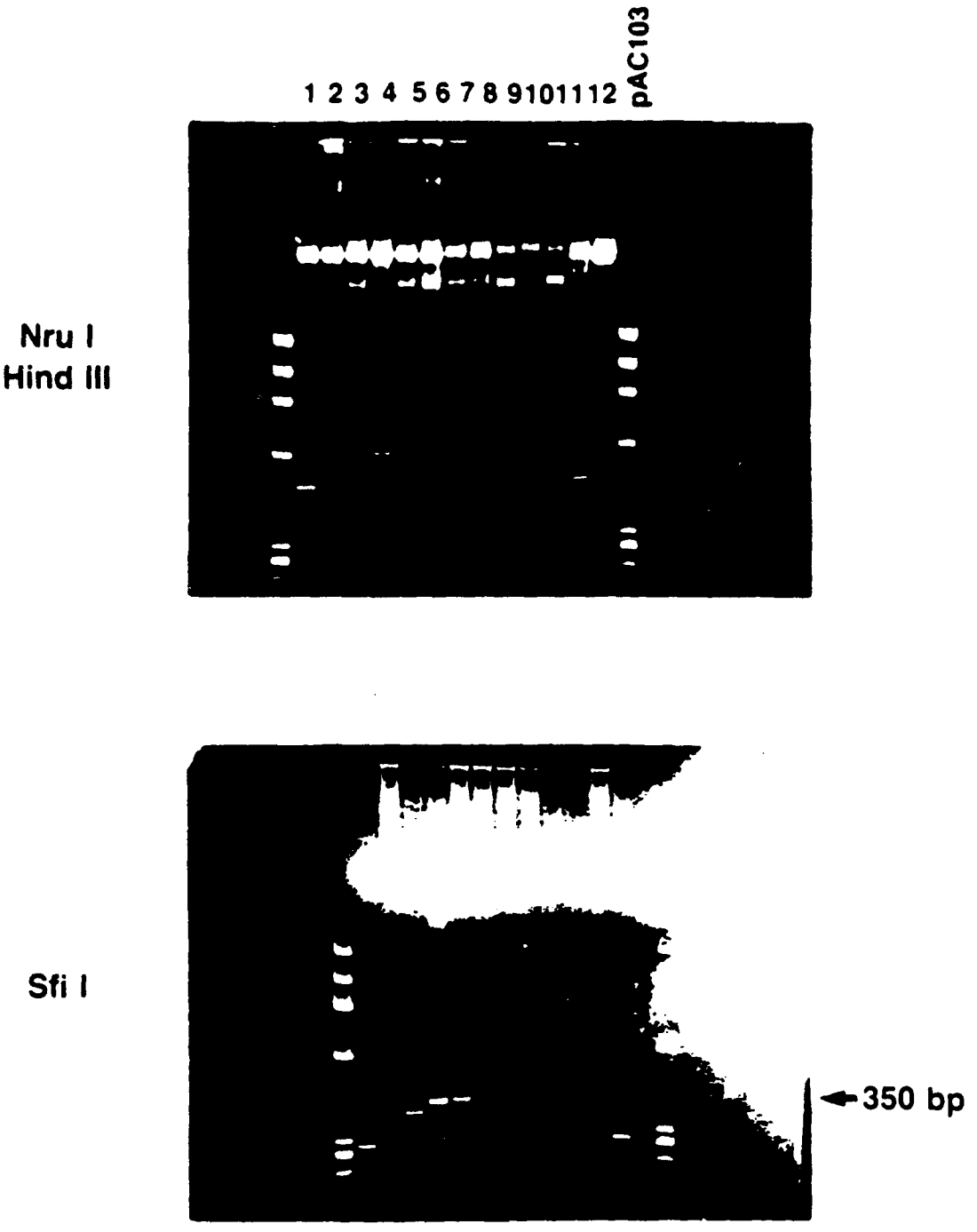


Figure 17. Deletion events occurring upon subcloning of a collagen analogue gene cassette. The 350 bp gene cassette from pAC103 is comprised of the 9 bp repeat 5'-GGTCCGCCG-3'. The cassette was gel purified and cloned into pAR using strain DG113d. Lanes 1-12 are plasmid DNA from 12 individual transformed colonies restricted with NruI+HindIII or with SfiI. Only isolate #9 appears to have the full length gene cassette. The remainder either have no cassette (#7-11), multiple cassettes (#2, 4 and 6), or single smaller cassettes (#1, 3, 5, and 12). Isolate #1 has two small cassettes cloned in tandem.

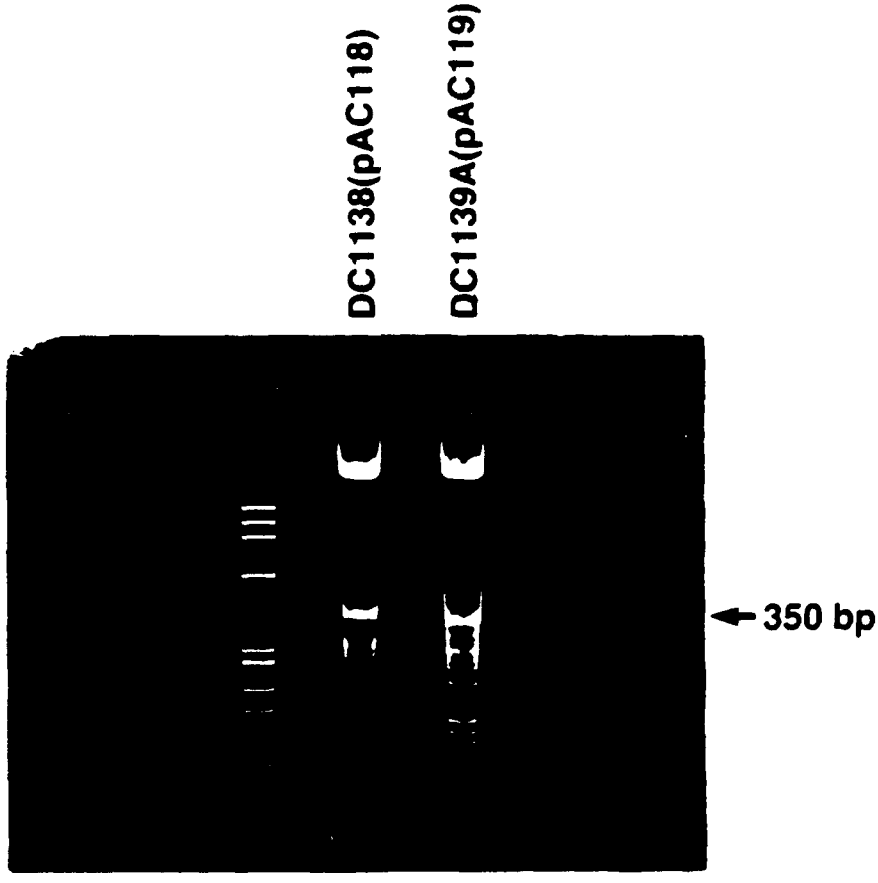


Figure 18. Distribution of collagen analogue gene cassette deletions. A 350 bp gene cassette was gel purified from pAC103, ligated into pAV6, and subcloned into hosts possessing (DC1138) or lacking (DC1139A) the functional (repressed) lambda *red* recombination system. Five ml overnight cultures were prepared and diluted 1/250 into 25 ml LB broth the following morning. Cultures were then expanded to 525 ml after the cells reached mid-log phase and were allowed to grow overnight. Plasmid was isolated by the alkaline lysis method [H.C. Birnboim and J. Doly, Nucleic Acids Res. 7, 1513-1523 (1979)]. Ten ug of each plasmid DNA was digested with NdeI+HindIII to liberate cassette DNA from the plasmid. In each case a ladder of bands is observed upon large-scale culture, and the difference in size corresponds to the 4 bp repeat in the DNA sequence.

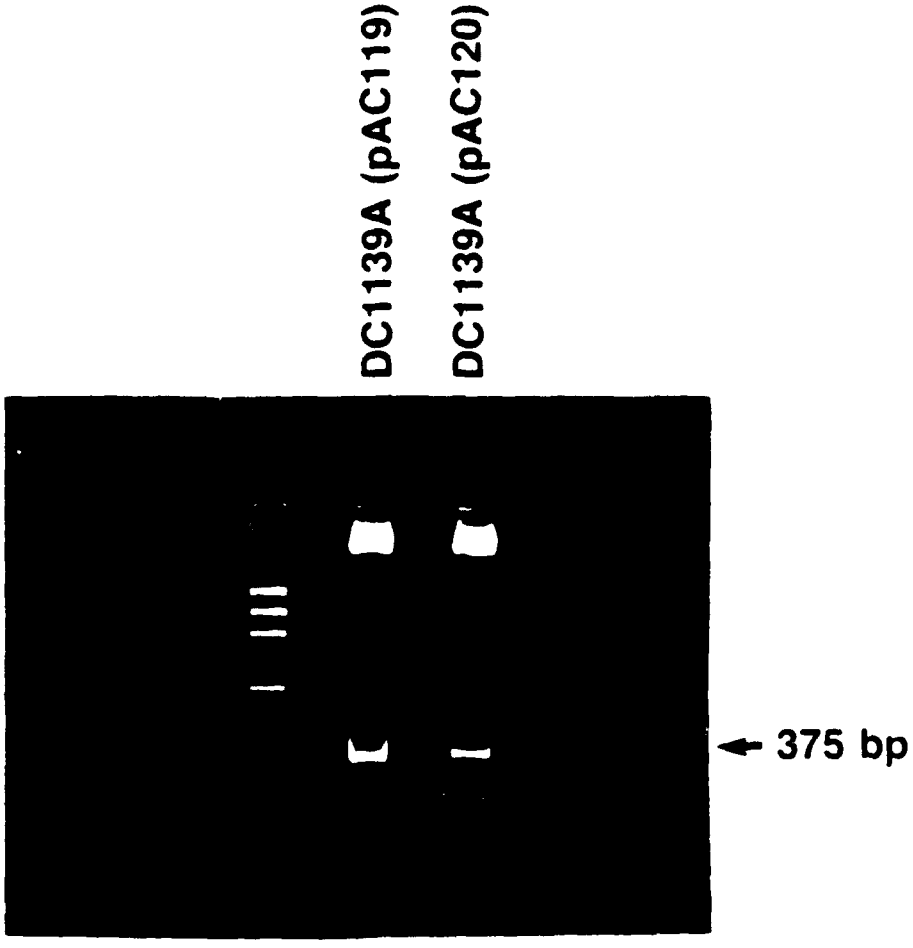


Figure 19. Comparison of deletions in collagen analogue gene cassettes containing 9 bp or 54 bp repeat units. The gel shows accumulation of deletions in collagen analogue gene cassettes upon large-scale culturing. pAC119 harbors a collagen analogue gene cassette made with a 9 bp repeat unit, while that contained in pAC120 is made with a 54 bp repeat. See Figure 18 for further experimental details.

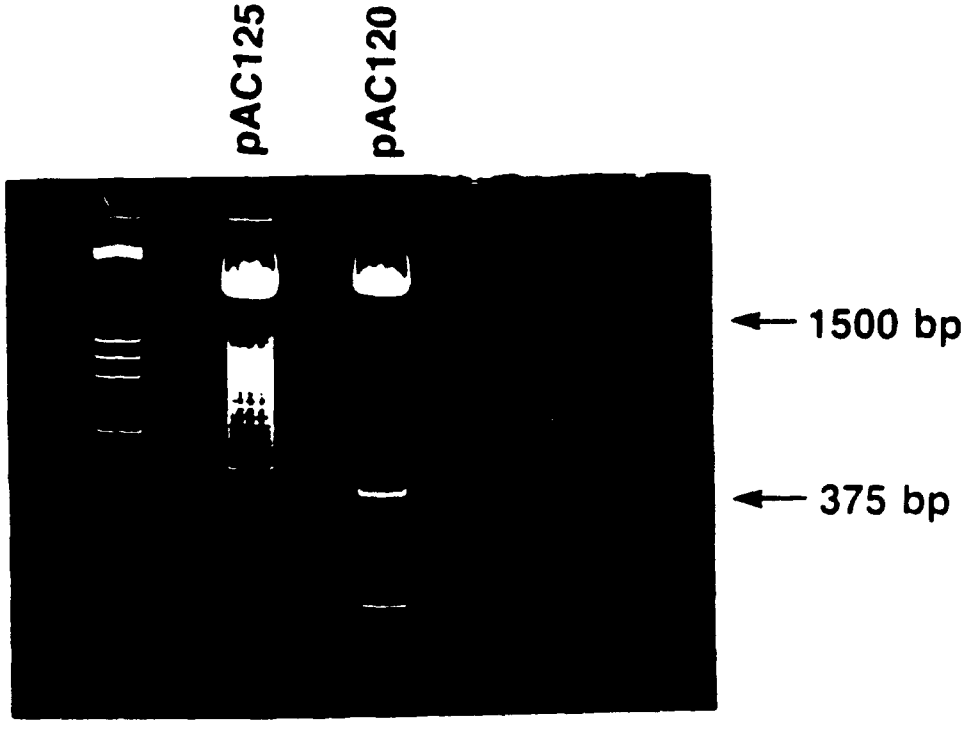


Figure 20. Stability of a large collagen analogue gene comprised of 4 tandem repeating gene cassettes. A 375 bp collagen analogue gene cassette was gel purified, ligated, and the 1.5 kb DNA fragment made from 4 tandem linked cassettes was again gel purified. This large DNA fragment was ligated into pAV6 and cloned into E. coli DC113e. The distribution of deletions in the large gene (pAC125) and the original cassette from which it is made (pAC120) is shown after large-scale culturing. Deletions corresponding to the 54 bp repeat unit of the original cassette predominate in the plasmid of each cassette. See Figure 19 for further experimental details.

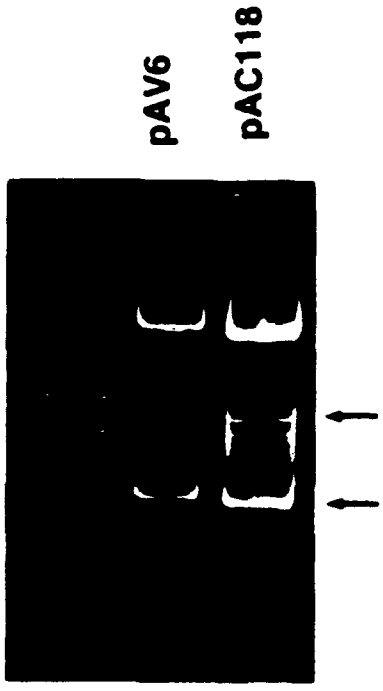


Figure 21. Complete deletion of the collagen analogue gene cassette from a majority of recombinant plasmids during culturing. Plasmid pAC118, containing a 300 bp collagen analogue gene cassette, or the progenitor plasmid pAV6 were isolated from large-scale cultures and 10 ml aliquots were digested with EcoRI+HindIII prior to 5% polyacrylamide gel analysis. The top arrow indicates the position of the DNA band containing the full-length collagen analogue gene cassette, while the bottom arrow highlights the deleted band which appears to contain only vector DNA. Plasmid was isolated from confluent cultures that had been diluted a total of about 1:50,000 from frozen glycerol stock cultures. The left lane contains ΦX174-HaeIII and lambda-HindIII DNA standards.

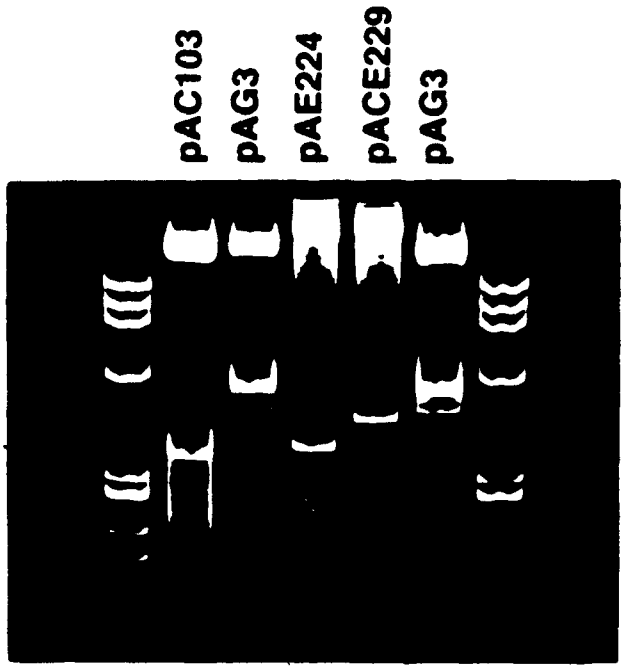


Figure 22. Comparative stability of collagen (pAC103), M. edulis polyphenolic protein (pAG3), and elastin (pAE224) analogue gene cassettes. Plasmid was isolated from host strain DC136 and analyzed as described in the legend to Figure 18, except for one pAG3 sample (third lane), which was isolated by gentle detergent lysis (G.N. Godson and D. Vapnek, Biochim. Biophys. Acta. 299, 516-520 (1973)). pACE229 contains a tandem cocassette comprised of a collagen and an elastin analogue. All plasmids show some accumulation of partially deleted cassettes, particularly collagen analogue gene cassette-containing plasmids and, to a lesser extent, pAG3. Plasmid pAE224 shows the least accumulation of partially deleted cassettes.

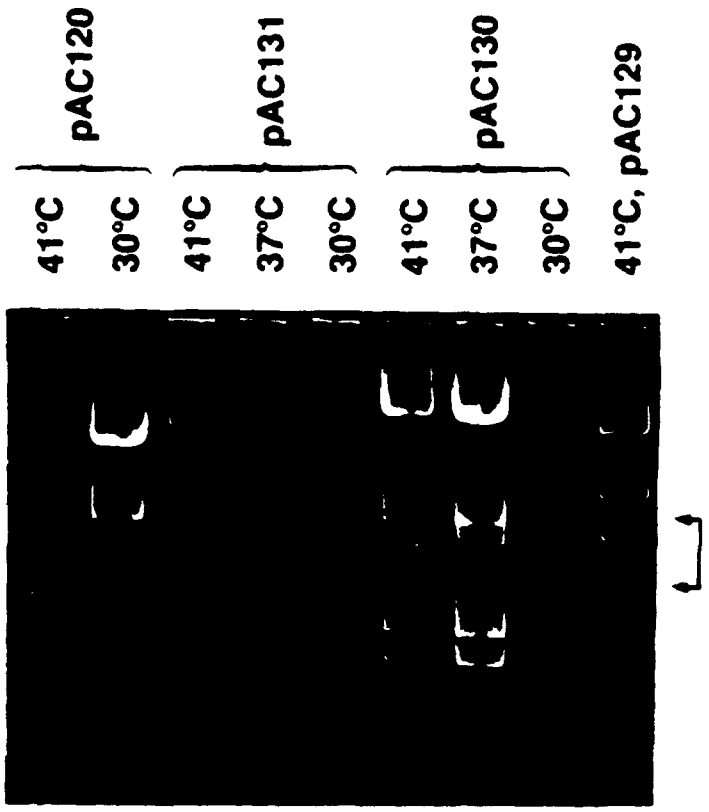


Figure 23. Collagen analogue gene cassette deletions as a function of cI857 repressor levels. The indicated plasmids were isolated from large-scale preparations and analyzed essentially as described in the legend to Figure 21 after growth of the cultures at the temperatures shown. The extra DNA bands are due to the presence of the DNA fragment carrying the cI857 gene. Large accumulations of partially deleted gene cassettes can be seen in the region between the two arrows of all lanes.