Journal of Supramolecular Structure
and Cellular Biochemistry 17:27-36 (1981)
Cellular Recognition 311-320

# Microchemical Instrumentation

**Leroy Hood, Michael Hunkapiller, Rodney Hewick, Charles E. Giffin, and William J. Dreyer**

*Division of Biology, California Institute of Technology, Pasadena, California 91125 (L.H., M.H., R.H., W.J.D.) and the Jet Propulsion Laboratory, Pasadena, California 91109 (C.E.G)*

A number of cell surface molecules of great theoretical and practical importance simply cannot be obtained in amounts sufficient for molecular analysis using conventional methods and instrumentation. Because of our interest in such studies, we began about eight years ago to explore the possibility of developing new instrumentation for the sequence analysis of very small quantities of polypeptide chains. These efforts have led to the development of two microsequenators which employ one thousandth to one ten-thousandth the material used in the original sequenator described by Per Edman. In addition, in conjunction with the explosion of recombinant DNA techniques, we also have begun to develop instrumentation for the sequence analysis and synthesis of DNA molecules. In this paper we describe briefly several new instruments that have been developed at Caltech. We believe this new instrumentation in conjunction with the recombinant DNA and hybridoma technologies will provide unique opportunities to analyze cell-surface molecules in the years ahead.

**Key words: Edman degradation, protein sequencing, DNA synthesis, peptide synthesis, mass spectrometer, genetic engineering**

## A SPINNING CUP SEQUENATOR

The sequence analysis of proteins and polypeptides is currently carried out using the Edman reaction [1]. This is a cyclic reaction in which under basic conditions the $\alpha$ amino group of the polypeptide chain is coupled to phenylisothiocyanate, and under acidic conditions the resulting derivatized amino acid is cleaved from the remainder of the polypeptide chain (Fig.1). The resulting amino acid derivatives (thiazolinones) are then converted to stable phenylthiohydantoin (PTH) amino acid derivatives. The PTH residue is routinely analyzed by high pressure liquid chromatography and the remaining polypeptide chain then is available for a second degradative reaction. Per Edman automated this chemical reaction in 1967 by developing the first protein sequenator (Fig. 2). This machine is basically a sophisticated plumbing device for carrying out the Edman chemistry. The protein is finely layered on the side of the spinning cup and reagents and solvents are delivered to this sample in a sequential fashion carrying out the Edman degradation which culminates in the cleavage of the N-terminal amino acid residue from the polypeptide chain and its delivery into a fraction collector for subsequent analysis. The original Edman sequenator required approximately 50–100 nmole of polypeptide for extended sequence analysis.

Received April 19, 1981; accepted April 21, 1981.

0275-3723/81/1701-0027$03.00 

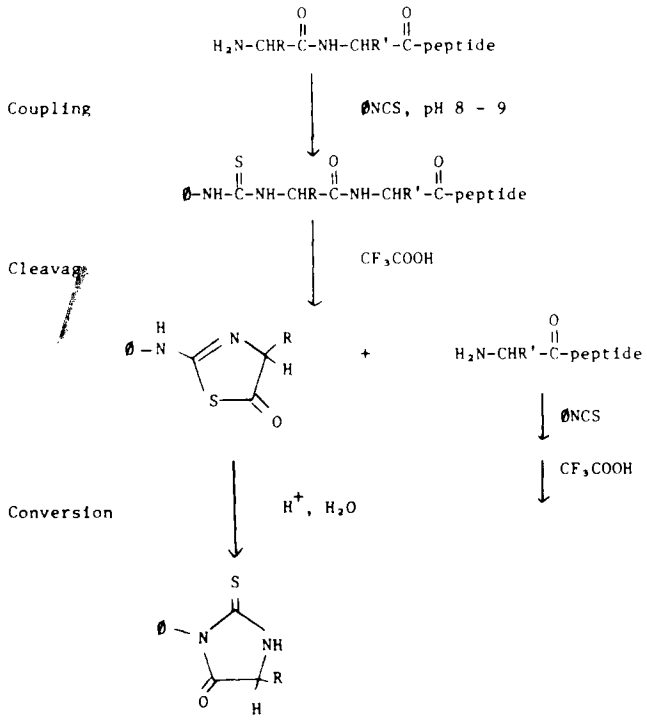


Fig. 1. Edman degradation.

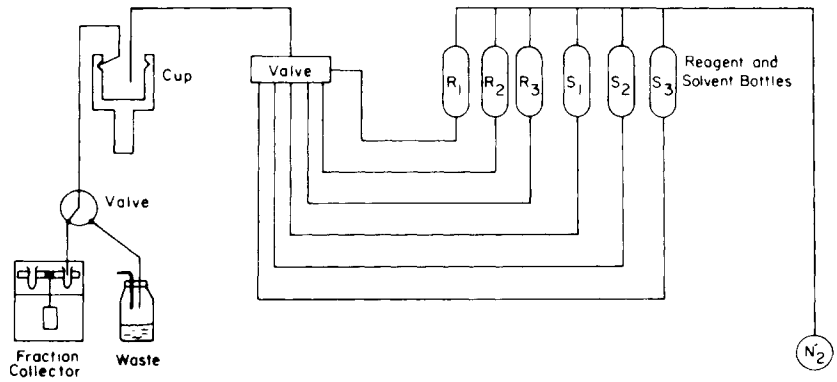


Fig. 2. Schematic diagram of spinning cup sequenator.

About four years ago we began to develop a new spinning cup sequenator with several important modifications designed to reduce the quantity of sample required [2]. These modifications included the development of a new type of valve, the introduction of a liquid nitrogen trap, the addition of a chamber for

automatically converting the thiazolinone derivative to its stable PTH amino acid form, the introduction of highly purified argon in the system to exclude side reactions caused by oxygen, the design of a new programmer that could be flexibly programmed to cleave amino acid derivatives under appropriate conditions (e.g., longer times for proline residues and shorter times for glutamine residues), and the use of a straight edge (rather than an undercut) cup as originally used by Edman and Begg [1] for optimal solvent washing of the polypeptide sample. The liquid nitrogen trap, valves, and automatic conversion chamber were modifications from improvements developed by Witmann-Liebold [3]. In addition, we exhaustively purified the solvents and reagents of the Edman reaction, made important modifications in the Edman reactions to optimize their efficiency, and developed the very sensitive, high pressure liquid chromatographic system for analysis of the PTH amino acid derivatives [4]. Moreover, we introduced a procedure whereby Polybrene could be used to facilitate the sequencing of microsamples; this turned out to be an important modification because this polymeric quaternary ammonium salt interacts with the carboxyl groups of polypeptide chains and the glass walls of the spinning cup to prevent polypeptides from being washed out of the spinning cup rapidly. With the Caltech spinning cup sequenator, we have been able to analyze proteins and polypeptides at the 20-picomole level. In Table I are given representative samples of the kinds of amino acid sequence analyses that we have been able to carry out on a great variety of polypeptide chains. Three points should be stressed. First, extended sequence analyses (greater than 25 residues) can be carried out on as little as 20 pmole of material. Second, short polypeptides can be sequenced in their entirety at reasonably low levels (e.g., dynorphin and the Aplysia neurohormones). Finally, very large polypeptides in the range of 200,000 daltons or more can be sequenced because relatively small quantities of material are required with the microtechniques described here. In conventional sequenators such large quantities of high molecular weight polypeptides are required for derivative detection that the proteins do not dissolve in the coupling buffer and even one or two steps of Edman analysis are difficult to carry out. Indeed, with this machine we have analyzed 122 of 166 residues in the human lymphoblastoid interferon using less than 40 $\mu$g of material. However, the Caltech spinning cup sequenator does have serious limitations both in the speed with which the cyclic Edman reaction can be carried out and its ultimate sensitivity in sequencing polypeptide chains. A number of years ago, with these limitations in mind, we also began to develop a more effective approach to the microsequence analysis of proteins. In this approach, a gas-liquid-solid-phase strategy was employed.

## GAS-LIQUID-SOLID-PHASE (GLS) MICROSEQUENATOR

In the GLS microsequenator, a spinning cup is replaced by a small column with an inert support—in our case fibrous glass [5] (Figs. 3, 4). The basic approach is that the polypeptide sample is embedded in a film of Polybrene coated on the surface of the glass fibers. The use of Polybrene eliminates the necessity of covalently attaching the sample to the support [6]. The coupling and cleavage reactions of the Edman chemistry are carried out using reagents in the gas phase, that is, dissolved in argon. The gas-phase reagents diffuse into the film of the

**Table I. Polypeptides Analyzed With the Caltech Spinning Cup Sequenator**

| Polypeptide | Residues identified | Quantity pmole | Quantity μg |
|---|---|---|---|
| Ribulose, 1-5, biphosphate carboxylase activation peptide | 17/17 | 1200 | 2.4 |
| Somatostatin from E coli | 14/14 | 1000 | 1.5 |
| Human carbonic anhydrase fragment | 23/23 | 200 | 0.5 |
| Aplysia egg laying hormone | 36/36 | 500 | 2.0 |
| Dynorphin (long leucine-enkephalin) | 17/17 | 500 | 1.0 |
| Porcine motilin | 22/22 | 800 | 2.0 |
| Porcine hypothalamic prosomatostatin | 28/28 | 600 | 2.0 |
| Sindbis virus E1 core lipopeptide | 31/38 | 250 | 1.0 |
| Goldfish scotophobin | 15/15 | 1000 | 1.7 |
| Sperm whale apomyoglobin | 36/153 | 20 | 0.3 |
| Human C'4 precursor | 25/1800 | 100 | 20 |
| Rat liver gap junction protein | 58/250 | 200 | 5.6 |
| Drosophila larval cuticle protein | 73/96 | 5000 | 50 |
| Human cytochrome oxidase subunit I | 20/540 | 20 | 1.3 |
| Torpedo acetylcholine receptor subunit III | 61/550 | 800 | 50 |
| Human fibroblast interferon | 24/166 | 18 | 0.4 |
| Human lymphoblastoid interferon | 31/166 | 75 | 1.4 |
| Human leukocyte interferon | 24/165 | 40 | 0.7 |
| Murine T cell growth factor | 31/230 | 40 | 1.0 |

Polybrene and polypeptide chain and carry out the Edman chemistry. The PTH amino acid derivative which is cleaved from the polypeptide chain is then washed from the solid support using an appropriate solvent in relatively small quantities (300 μl). This approach has several advantages. First, the Edman chemistry can be carried out at the rate of one cycle per 50 minutes. This allows one to analyze 28 residues a day as compared with 12–14 residues a day for the spinning cup sequenator. Second, the instrument requires 1/20 as much reagents and solvents. Hence the operation of this machine is far less expensive than its spinning cup counterpart. Third, this machine permits us to analyze model polypeptides at the 5-picomole level. Moreover, the Edman chemistry seems to be very efficient in that samples analyzed at the 10-nanomole level have repetitive yields of up to 98%. Thus we have been able to analyze 10-nmole samples out to 90 residues and 5-pmole samples for 25 or more residues. Table II shows the striking increase in sensitivity that has been provided by the sequenators developed at Caltech. Indeed, it seems theoretically possible to increase the sensitivity to the 1-picomole level which would be 0.02 μg of a model 15,000 molecular weight protein (myoglobin). When we can analyze samples at the 1-picomole level, we can turn the most sensitive analytic technique in molecular biology, two-dimensional gel electrophoresis, into a preparative method for analyzing polypeptide chains. We expect to be able to take a two-dimensional gel and isolate from that a single polypeptide spot and obtain detailed sequence information on the corresponding polypeptide.

We anticipated that the increased sensitivity of the GLS microsequenator would enable us to operate with sample amounts below the limits of detection for

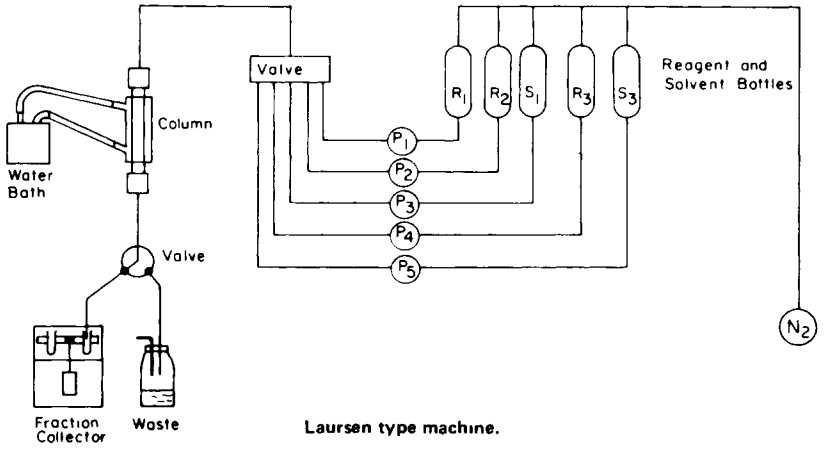


Fig. 3. Schematic diagram of Laursen solid phase sequenator.

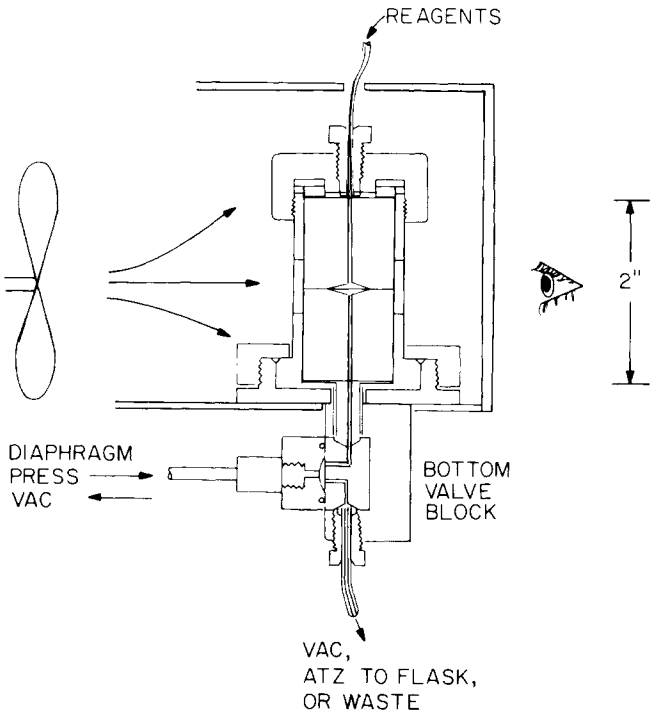


Fig. 4. Cartridge and exit delivery valve design for gas-phase solid state sequenator.

the high-pressure-liquid-chromatographic system. For this reason we have spent the past five or six years in collaboration with researchers at the Jet Propulsion Laboratory developing a new analytic instrument for the analysis of very small amounts of PTH derivatives—a highly sensitive mass spectrometer.

**Table II. Increase in Sensitivity of Automated Edman Degradation of Proteins***

| nmole | μg | Instrument |
|---|---|---|
| 100 | 4000 | 1967—Edman/Begg spinning cup |
| 10 | 400 | 1971—commercial spinning cup |
| 0.1 | 4 | 1978—CIT-modified commercial spinning cup |
| 0.02 | 0.8 | 1979—CIT microsequencing spinning cup |
| 0.005 | 0.2 | 1980—CIT gas/liquid solid phase |

*Amount of protein required for 20-residue N-terminal analysis.

## MINIATURIZED ELECTRO-OPTICAL MASS SPECTROMETER

A Caltech group, in collaboration with engineers at the Jet Propulsion Laboratory in Pasadena, have conceived and developed a new type of mass spectrometer [7–9] (Fig. 5). This mass spectrometer is small and stable and will hopefully be capable of being operated by a relatively inexperienced operator. The key feature regarding the development of this new mass spectrometer was the development of an electro-optical ion detector array which has the capacity to detect a single molecular ion (Fig. 6). The current design electro-optical array has 25,000 electron-multiplier detector units per cm length along the mass spectrometer focal plane (12.7 cm long). Each ion that impinges on one of the detector elements is converted to a pulse of 10,000 electrons. These electrons are then accelerated to a phosphor screen where a pulse of $10^6$ photons is generated and illuminates through fiber optics a 25 μm wide photodiode (5120 photodiodes view the focal plane covering a mass range of 25–500 amu). The accumulated charge across a capacitor on each photodiode becomes a measure of the number of ions that have illuminated that fraction of the mass spectrometer focal plane. Thus a mass spectrum can be simultaneously integrated by greater than 5000 separate photodiode detectors during a time frame of tens of milliseconds to many seconds depending on ion signal intensity. The output of the electro-optical detector is ultimately fed to a computer where the data are calibrated and analyzed. This mass spectrometer can analyze PTH amino acid derivatives in the femtogram range, thus making it greater than 1000-fold more sensitive than present HPLC instrumentation. Moreover, with the implementation of an appropriate sample injection system and computerized data processing, it should be possible to have a sample turnaround time of 5 minutes on the mass spectrometer. Thus we envision in the future multiplexing two or more GLS microsequenators to the mass spectrometer and, accordingly, being capable of analyzing up to 100 residues or so a day. The development of an automated sample injection system is now under way, and hopefully we will have this instrument finished within the next year and a half or so.

## DNA SYNTHESIZER

Dr. Marvin Caruthers at the University of Colorado has developed a solid phase synthetic technique for the synthesis of DNA oligomers [10]. This chemical synthetic procedure lends itself very well to automation using modular subcomponents of the GLS microsequenator described earlier. Indeed, we have just finished construction in our laboratory of a prototype solid phase DNA synthesizer and are now in the process of testing it (Fig. 7). From the manual DNA

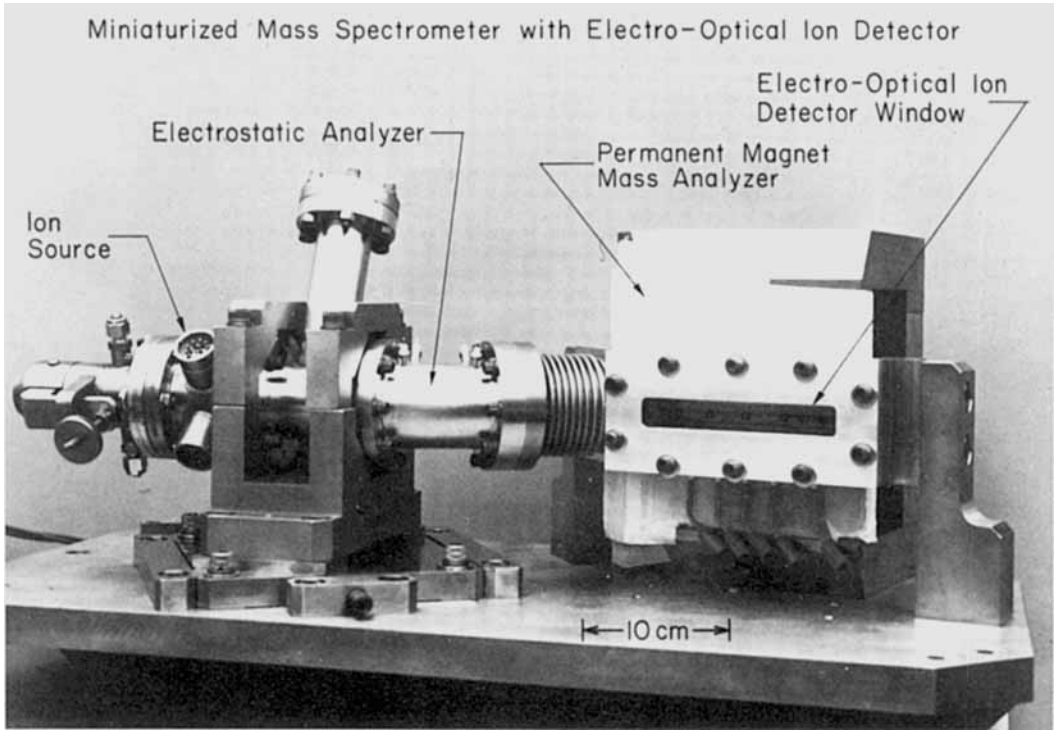


Fig 5 The miniaturized electro-optical mass spectrometer

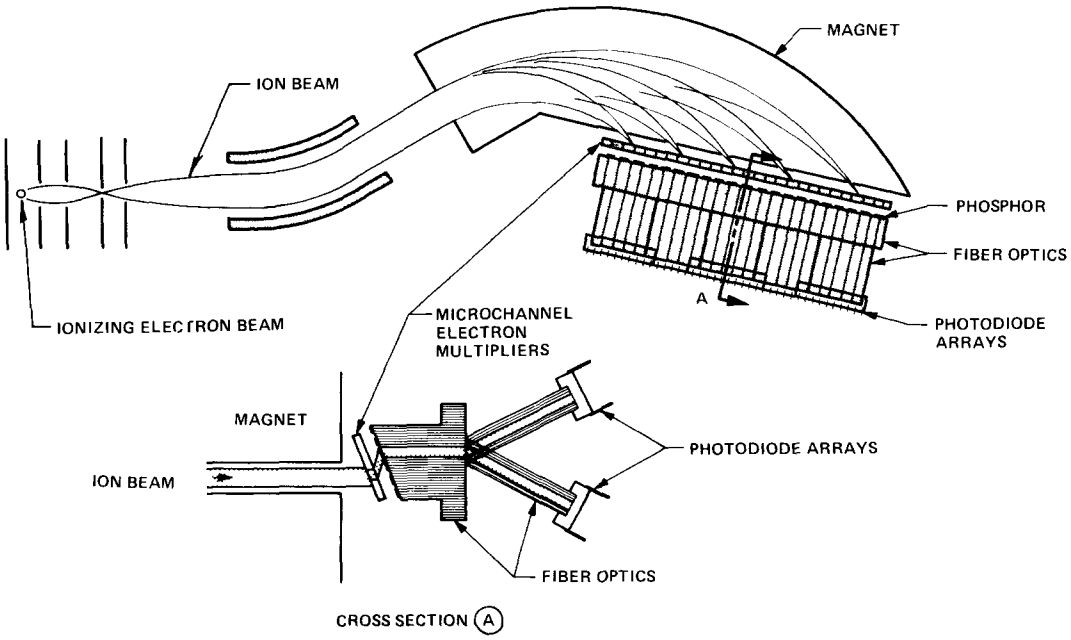


Fig 6 Schematic diagram of mass spectrometer with electro-optical ion detector array

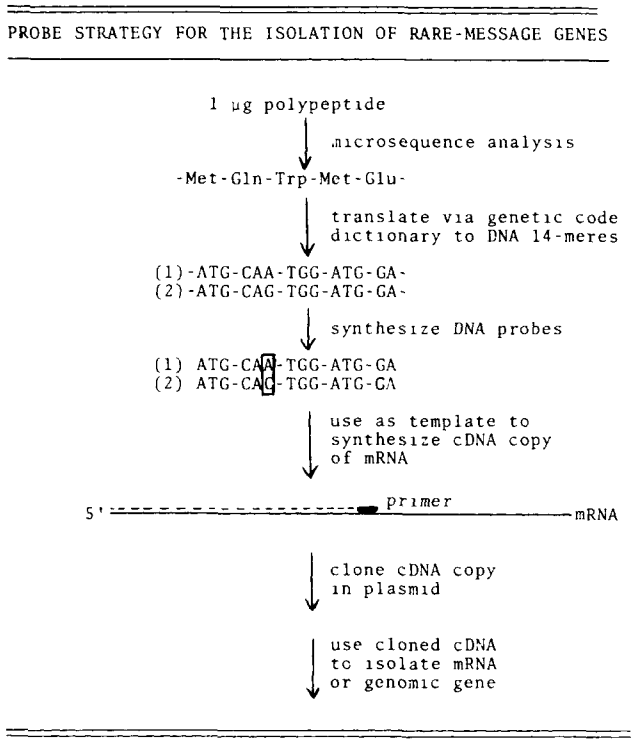


Fig. 7. Schematic diagram of prototype DNA synthesizer.

synthetic chemistry that has been carried out to date, it appears that the repetitive yield for this DNA synthetic procedure is of the order of 95% or greater for each of the four DNA bases and that a cycle time should be complete in a half hour or less. Thus we should have the capacity to synthesize approximately 50 nucleotide linkages per day. Once again, extrapolating from the current manual chemistry it appears that this machine should easily be capable of synthesizing oligomers that are anywhere between 20 and 40–50 nucleotides in length. Accordingly, the DNA synthesizer will have a wide variety of uses in recombinant DNA technologies and these are listed in Table III.

## A NEW APPROACH TO THE ISOLATION OF RARE MESSAGE GENES

The combination of a gas-phase solid state sequenator with eventual 1-picomole sensitivity and an automated DNA synthesizer raises the possibility for a new approach to the isolation of rare message genes [11] (Fig. 8). Perhaps a tenth of a microgram of the protein can be isolated and analyzed using the automatic sequenator. Then the 30–50 residues so analyzed can be examined for stretches of four or five residues that have minimal ambiguity when translated by the genetic code dictionary back into DNA language. If this stretch of protein sequence has a twofold degeneracy as illustrated in Figure 8, then the 14-meroligo-

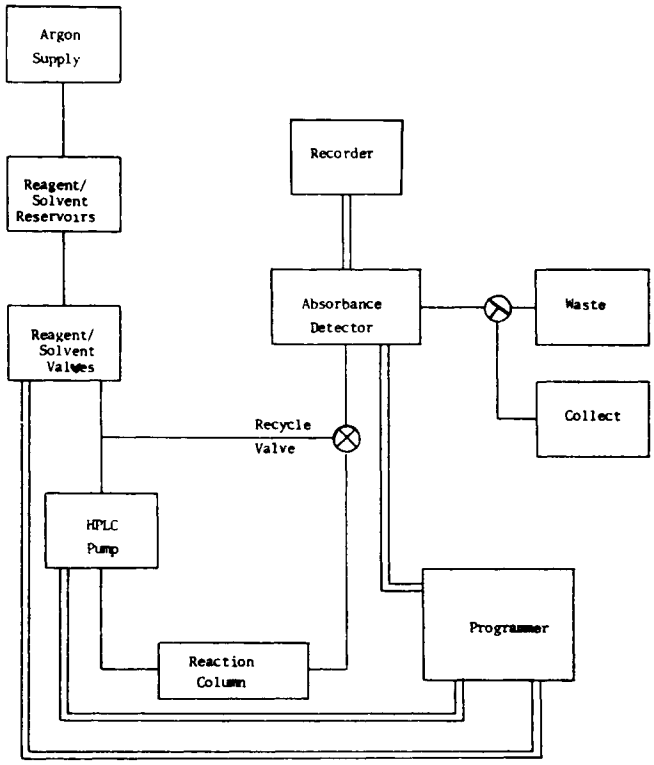


Fig. 8. Strategy for isolation of rare message genes based upon protein sequence analysis and synthesis of DNA primers.

**Table III. Uses of DNA Synthesizer**

1. Probes for isolation of rare-message genes
2. Synthesis of complete genes
3. In vitro mutagenesis
4. Tools for recombinant DNA procedures, e.g., linkers, etc.

nucleotides would be hybridized to complementary nucleotide sequences on the messenger RNA and there could serve as primers to initiate cDNA synthesis in the 5′ direction with reverse transcriptase. In this fashion the corresponding DNA segment could be cloned using conventional recombinant DNA techniques and, if long enough, it could be used to examine directly a genomic library for the corresponding gene. If the cDNA clone were not of sufficient length, it could be used to purify specific messenger RNA from which a full length cDNA clone for the corresponding gene could then be derived. This technique allows one to move directly from protein sequence to gene isolation and should be particularly

beneficial in the analysis of a variety of different types of rare message genes. Indeed this technique permits one to use in a very effective fashion three new technologies that are revolutionizing cell biology—the hybridoma technique [12] which generates monoclonal antibodies, our microinstrumentation facility, and the newly developed recombinant DNA technique. Clearly this type of approach is going to make it possible to isolate and characterize in detail a variety of different recognition molecules present on membranes that may be related to cell-cell recognition and to fundamental processes of differentiation in complex eukaryotic systems.

## ACKNOWLEDGMENTS

The work was supported by NIH grant GM 06965 and NSF grant PCM 80-05599, as well as a grant from the Weingart Foundation.

## REFERENCES

1. Edman P, Begg G: Eur J Biochem 1:80, 1967.
2. Hunkapiller MW, Hood LE: Science 207:523, 1980.
3. Wittman-Liebold B: Hoppe-Seyler's Z Physiol Chem 354:1415, 1973.
4. Johnson N, Hunkapiller MW, Hood LE: Anal Biochem 100:335, 1979.
5. Hewick R, Hunkapiller M, Hood L, Dreyer W: J Biol Chem 256:799, 1981.
6. Laursen RA: Eur J Biochem 20:89, 1971.
7. Dreyer WJ, Kuppermann A, Boettger HG, Giffin CE, Norris DD, Grotch SL, Theard LP: Clin Chem 20:998, 1974.
8. Giffin CI, Norris DD, Boettger HG: Mass Spectrometry and Ions Physics 15, 1975.
9. Boettger HG, Giffin CE, Norris DD, Dreyer WJ: In Talmi Y (ed): "Monograph on Image Detectors in Chemistry." American Chemical Society Press, 1979.
10. Stinson SC: Chem Eng News, February 2, 1981, p 17.
11. Noyes B, Mevarech M, Stein R, Agarwal K: Proc Natl Acad Sci USA 76:1770, 1979.
12. Milstein C: Sci Am, October 1980, p 66.