Chapter 3
25 Years of HIV-1 Biochemistry

David E. Ott

Retroviral biochemistry came to the forefront in the late 1960s, the co-discovery of reverse transcriptase by Howard Temin's and David Baltimore's laboratories [1, 2]. Retroviruses, originally referred to as RNA tumor viruses, oncoviruses, or oncornaviruses, were being studied for their induction of cancer mostly in avian and murine systems [3, 4]. The term retrovirus first appeared in 1975 [5], reflecting the realization that these RNA viruses go against the "central dogma" of molecular biology, which held that DNA was the principal source of cellular information with genes being expressed by transcription into RNA which is then translated into proteins [6, 7]. RNA tumor viruses were found to reverse this going from RNA to a viral DNA form that is integrated into the host to form a provirus, a mechanism first hypothesized in the 1960s by Howard Temin (see his Nobel Prize lecture [8] for a historical review). The HIV-1 proviral form acts as a stable locus of genes in the cell, producing both viral proteins and the viral RNA genome through a complex series of molecular gymnastics. Hence, the term retrovirus for backward virus was coined. However, it is important to remember that this major advance as well as other in retroviral biochemistry was before the advent of many of the modern biochemical tools, instead relying on rudimentary forms of analysis, protein purification, enzymology, chromatography, protein sequencing [9, 10], spectrometry, velocity sedimentation, density centrifugation, and immunodiffusion. However, progress was slow and the ability to accurately analyze complex mixtures of proteins was nonexistent, unlike today with current mass spectrometry methods.

As discussed in chapter "Introduction: HIV-1 Proteomics, Why Should One Care?", the study of retroviruses has produced an abundant harvest of insights into cancer, immunology, cell biology, antiviral vaccines, biochemistry, and genetics.

D.E. Ott, Ph.D. (✉)
Leidos Biomedical Research, Inc., Frederick National Laboratory
for Cancer Research, 1050 Boyles Street, Frederick, MD 21702, USA
e-mail: ottde@mail.nih.gov

D.R.M. Graham, D.E. Ott (eds.), *HIV-1 Proteomics*,
DOI 10.1007/978-1-4939-6542-7_3

The mechanistic knowledge gleaned from studying the prototypic avian and murine retroviruses laid an essential foundation for the rapid progress in characterizing HIV-1 and AIDS drugs. The early, pre-HIV, study of retroviruses was carried out with fairly basic techniques, basic chromatography, and enzyme assays. Truly these early retrovirologists accomplished amazing things, relying on their brains and careful, well-thought-out experiments more than high-tech techniques. This chapter provides an introduction and discussion of the more modern "classical" techniques which complement the newer "proteomic" mass spectrometry approach and are as essential today as they were decades ago.

Aspects of HIV-1 Particle Biochemistry

Classical biochemistry starts with examining a highly complex mixture of material usually in the form of a tissue from which cells are isolated and lysed, and the components, protein, lipid, or carbohydrate, of interest, are purified and studied. For proteins, this is typically an involved process to isolate the relatively tiny amounts of the protein of interest in the vast sea of cellular proteins. Compared to classical cellular biochemistry, retroviral protein biochemistry is easier since the virus does most of the purification work by releasing particles into the cell culture medium leaving the complexity of the cell behind. By doing so, viruses conveniently purify themselves for the researcher, who can straightforwardly isolate HIV-1 virions using their biophysical properties, either by their density or their size using density or sedimentation centrifugation.

Within the last 15 years, there has been a strong interest in the cellular proteins incorporated into HIV-1 virions [11–13]. Viruses, by definition, require cells to replicate. HIV-1, retroviruses in general, has a relatively small genome when compared to the large DNA viruses which have extensive genomes that code for many proteins which provide independent viral replication functions and modulation cell functions for immune escape and manipulation of the cell for the advantage of the virus. Therefore HIV-1 relies on mostly cellular proteins to replicate. One way to study this is to examine the host proteins found in virions (see chapter "HIV-1 Biology at the Protein Level" for an extended discussion). Host proteins can be incorporated into virions by just being present at the site of HIV-1 budding, being taken up as bystanders due to their presence in the plasma membrane as the particle buds through the membrane. These proteins, while not specifically incorporated, still provide for clues to the site of budding and cell type producing the virus. Proteins could also be incorporated as partners when they interact with viral proteins as well as actively assist in assembly and virion production. Also, HIV-1 could incorporate cellular proteins as captives to assist in post-assembly functions such as immune evasion and promotion of cellular infection.

Thus, while the study of the cellular proteins in virions is important, the potential for contamination is a critical concern because while the origin of the HIV-1 proteins in virions is obvious, great care must be carried out to selectively detect

those cellular proteins both in and on the virion versus those that are present as contaminants. The essential question is how to show that a protein is on or in the virion as opposed to being merely being in "purified" virus preparations. For instance, one classical method for surface proteins is to specifically immunoprecipitate the virus and then assay for a viral protein, typically capsid. This is essentially a qualitative method and unable to examine proteins inside the virion. For quantitative studies, the virions must be purified and the particles examined directly. While separation of virions from culture media and most proteins released by cells can be accomplished by using centrifugation, still a significant amount of cellular proteins co-purify with retroviral virions, being present in vesicles (microvesicles or exosomes [12, 14]) that have the same size and density as virions. Thus, studies seeking to identify and characterize cellular proteins in the HIV-1 particles require high levels of purification with strict controls that demonstrate that the protein(s) of interest are removed. Two techniques, protease digestion of virions and CD45 immunoaffinity depletion, are discussed in chapter 2: "HIV-1 Biology at the Protein Level" so they will not be recapitulated here. Nevertheless, in biochemistry, either classical or high-tech mass spectrometry, the purity of the sample is critical to draw accurate conclusions from the results.

HIV-1 Proteomics Before There Was Proteomics

Before the current mass spectrometric sequencing capabilities, analysis of viral and cellular proteins in retroviruses was done with well-developed techniques that relied on classical biochemistry which lack the throughput of proteomics but still identified and characterized the HIV-1 viral proteins and their modifications as wells as many cellular proteins present both in and on HIV-1 particles. In the days before proteomics, analysis of proteins in the virion took a brute-force approach (techniques are discussed below): isolate the various proteins in the virion by chromatography, run the fractions on a protein gel, blot the proteins on a gel, and cut out bands for automated microsequencing. This approach was labor intensive and took months/years to complete. In contrast, this can be done in a couple of days with much higher sensitivity with current MS/MS spectrometry. Yet the classical approach and its associated techniques still have some advantages. The following is a review of these classical biochemistry techniques, their strengths and limitations.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

An essential advance in biochemistry arrived with the advent of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), a critical technique that is widely taken for granted today. Before SDS-PAGE proteins were electrophoresed on a variety of gel substrates and buffer systems. Because proteins can have neutral,

positive, or negative charges, the samples were loaded in the center of the gel and electrophoresed with proteins migrating according to their charge, positive to the cathode side and negative to the anode. Thus, while proteins could be separated using this system, relative size was not determined. Note that this was not isoelectric focusing which yields the protein's isoelectric point, a measure of the protein's pH properties [15]. This changed in 1966 with the first use of the negatively charged detergent sodium dodecyl sulfate (SDS) in the sample which was electrophoresed in thin polyacrylamide tube gels (see references in [16]). Since the SDS not only solubilizes most proteins, but imparts a net negative charge to the protein, the proteins migrate in only one direction, toward the anode. Thus, proteins could be loaded on the top of the gel and unidirectionally electrophoresed down the gel. It was soon after the introduction of SDS-PAGE (1967) that it was recognized that the SDS bound most proteins according to their mass. Thus, amount of negative charge imparted to the proteins by SDS is, for the most part, proportional to the molecular mass of the protein [16], a major breakthrough. However, at the time, polyacrylamide was cast in tube gels which meant every sample had to be run in an independent tube, making comparison of samples difficult. Now samples with different proteins could be meaningfully compared based on a consistent measurable property, their mass. The final major step was the development of the slab-gel SDS-PAGE system by Laemmli in 1970 [17] which involved casting a polyacrylamide gel as a rectangular slab using a Trizma HCl (Tris)-glycine-SDS buffer system and a discontinuous gel system that used a small "stacking" gel on top the analytical gel to focus the protein sample before it entered the main gel. This system now allowed a wide variety of protein samples to be resolved side by side with a molecular mass marker for accurate comparison and determination of their apparent molecular weight. While later innovations such as gels with increasing gel concentrations, gradient gels, and alternative buffer systems have appeared, the Laemmli system remains the basis for analyzing proteins, and the Laemmli paper is one of the most cited in biological research [17].

Immunoblotting/Western Blotting

The discovery of the DNA code opened a host of new ways to analyze DNA, a molecule once thought of as a gooey, stringy nuisance contaminant by protein biochemists. One of these was the Southern blot, eponymously named for its inventor presented in another seminal paper [18, 19]. For this technique, DNA was separated on an agarose gel, blotted onto a sheet on nitrocellulose, and then hybridized with specific radiolabeled DNA probes to reveal the presence of the specific DNA in the mixture. This watershed advance in nucleic acid research was soon followed by applying this technique to the analysis of RNA (Northern) and protein (Western blot or immunoblot). For the Western blot, SDS-PAGE gels are electroblotted onto polyvinylidene fluoride (PVDF) membranes and reacted with serum or antibodies specific for a protein or proteins. The bound antibodies are visualized by various means, having been labeled directly with a fluorophore, enzyme, or radioisotope before the

analysis or detected secondarily using a labeled anti-immunoglobulin antibody which binds to the primary antibody. This technique brought even more analytical power to SDS-PAGE techniques by being able to identify specific proteins as immunoreactive bands in the sample lane, no matter how complex the mixture.

Western blots are an essential tool for the analysis of proteins and are widely used by HIV researchers. There are primarily two types of immune reagents used to visualize proteins of interest, antiserum or monoclonal antibody. The key to the success of this technique is the strength and specificity of these reagents, each of which has its distinct advantages. Antiserum (polyclonal sera) made to the whole or just a portion of the protein of interest commonly recognizes more than one epitope and thus can detect more than one portion of the protein. Also, these immune reagents can be stronger due to dominant epitope specificities in serum. For instance, we have some antisera that readily detect viral proteins at a 1:500,000 dilution for analysis compared to a typical dilution of 1:4000 for antisera and 1:1000 for monoclonal antibodies. However, this is not always the case as there are plenty of weak antisera and strong monoclonal antibody titers. Additionally, because there can be an array of antigen-specific antibodies present, antisera can recognize several epitopes within a protein; Western blots using antisera are much more plastic in detecting proteins. Therefore, variant proteins such as mutants, isotypes, splice variants, related proteins, and multiple fragments of proteolytically processed proteins may be detected. Antiserum made to short defined peptides do not have these advantages since they typically elicit only one epitope specificity. A disadvantage to antisera is that it contains not only the desired antibodies, but the other antibodies in present in the serum, so it is possible that these reagents could produce higher background noise and nonspecific bands. For instance, many antisera are produced by linking a peptide to bovine serum albumin which is also a component of cell culture media and commonly used in molecular mass standards. Thus, these types of peptide antisera can lead to strong unwanted bands in blots that can overwhelm and obscure bands from the desired proteins. Finally, the supply of any given antiserum is finite and cannot be precisely reproduced.

Monoclonal antibodies recognize only one epitope, thus only one part of the protein. Because the epitope is usually defined, being made to a peptide, the interpretation of the results is more straightforward. This can be useful for mapping proteolytic fragments or splice variants with two or more antibodies. Also, a monoclonal antibody can have less background, having only one antibody present instead of the many specific and much more nonspecific antibodies present in the serum. Despite having only one specificity, there are many monoclonal antibodies that detect one or more irrelevant bands, due to the antibody cross-reacting with similar epitope on another protein. Finally, monoclonal antibodies are produced from hybridoma cell lines either in vitro or in vivo, so a good antibody can be reproduced. Also, due to their purity, monoclonal antibodies are commonly directly labeled for detection of proteins and as secondary detection reagents in Western blots.

Despite the analytical power and sensitivity of current mass spectrometry for analyzing HIV-1, both SDS-PAGE and its logical extension, the Western blot, still complement modern "proteomic" techniques. First, they are rapid and easily interpretable, with some procedures giving results within a couple of hours with low cost. The

instrumentation is comparatively simple so each laboratory member can have their own equipment and carry out analyses independently. Most importantly, these techniques examine the whole protein, not just the sequence of peptide fragments by MS/MS spectrometry in complex mixtures of proteins which leads to ambiguity: was this sequence from a pro-form of a protein, a processed active form, or a variant, or, in the case of the Gag polyprotein, is the peptide from Gag or just one of its mature proteins? Thus, many researchers confirm proteomic protein identifications with a Western blot. Alternatively, the complexity of a mass spectrometry sequencing sample can be reduced by fractionating by size. The sample can be separated on an SDS-PAGE gel which is the cut into size fractions and the proteins in each slice digested and analyzed to provide the sequences present in a specific molecular mass range. An Achilles heel of mass spectrometry is the bioinformatics side where the peptide sequences are matched to proteins. Because peptide sequences are compared to expected values in databases, genetic polymorphisms, unexpected protein modifications, and other factors can cause vital peptide to be missed obscuring the identification of proteins. Finally, these classic biochemistry and proteomic techniques have opposing, thus complementary natures. Western blots look inward: one needs to decide what protein to interrogate the blot for a principally yes or no answer. Since these antibodies need to be chosen, one needs to test from a hypothesis of what protein is of interest as it is obviously not profitable to guess with tens of thousands of proteins in a cell. In contrast, MS/MS spectrometry sequencing is outward looking, one harvests information on nearly all of the proteins in a complex mixture simultaneously, yet one does not study the intact protein. Mass spectrometry and Western blots work hand in hand, proteomics indentifying candidates and blots providing valuable information about the form of the protein and its posttranslational modifications.

Reversed-Phase High-Pressure Chromatography

Chromatography in many different forms has been used for over a century to separate materials. Biochemically, paper, silica plates, and columns (mostly liquid in biology) with various packing materials have been used for years to separate tissue and cell extracts into pure proteins. Four principle types of liquid chromatography column are molecular sieve/gel exclusion which separates on the basis of size; ion exchange which exploits differences in charge at a pH, i.e., their isoelectric point (pI); affinity which relies on binding of a protein to a material, DNA, cofactor, protein, or specific antibody; and reversed phase which separates proteins based on the hydrophobic character. This last method employs a hydrophobic column material through which sample in a hydrophilic solvent is pumped. Since all proteins have some hydrophobic character, most bind to the column packing matrix. The proteins are then separated by increasing the hydrophobicity of the solvent flowing over the column in a gradient. Proteins are eluted from the nonpolar matrix when they partition into the solvent according to their hydrophobicity, i.e., they are more soluble in the solvent than on the matrix. This process is carried out at high pressure to produce a rapid flow rate that is precisely

controlled so that the proteins eluted off the column are resolved in narrow peaks at distinct and reproducible times. The primary benefits of reversed-phase high-pressure liquid chromatography (r-pHPLC) are high resolution, speed, and reproducibility; many virion proteins can be purified in one column run. The r-pHPLC technique is better than ionic exchange because the hydrophobic nature of a protein is typically little changed, whereas a protein charge can vary widely based on oxidation, reduction, and differential posttranslational modification of the proteins. Molecular exclusion columns have relatively low-resolution power compared to r-HPLC so they can rarely resolve proteins out of complex mixtures. One of the early downsides to using r-HPLC is that the column sizes were relatively large, therefore requiring large amounts of sample. This is good for preparative work, for separating proteins from milligrams of purified virus, concentrated from liters of virus production, but not for the analytical microgram to tens of nanogram amounts produced from cell transfections or material isolated directly from primary sources. The advent of microbore HPLC techniques, using ~2 mm columns with single digit micron-sized column material, now allows for small amounts of virus (from >1 to 0.1 μg) to be rapidly analyzed [20]. With high precision pumps and programmable gradient controllers, r-HPLC can resolve both the viral and host proteins in the virion at reproducible times allowing for matching of chromatograms by overlaying the profiles to compare several samples. Also, when using a UV detector in the near UV range (206 nm), where the absorbance of amino acids is the same, the peak area is directly proportional to mass, allowing for comparative measurements between protein peaks in the chromatograph.

The downsides to r-HPLC are that this technique requires that the proteins be denatured. Thus, the isolation of active enzymes or intact protein complexes such as multiprotein complexes is not possible, unless they can be refolded/reassociated post-purification. Another drawback is that r-HPLC samples cannot contain detergent, a common technique for producing cell extracts, as this binds to the column and interferes with the partitioning of the solute proteins between the column and the solvent. While both of these limit r-HPLC utility for analysis of cells, the goal of analyzing retrovirus particles is to examine their protein composition rather than enzymology. Also since the complexity of proteins in virions is relatively simple compared to cells, they can be lysed directly in the r-HPLC loading buffer without extraction and isolation procedures involving detergent. Despite these limitations, r-pHPLC remains a powerful technique for the examination of both viral and cellular proteins in HIV-1, simultaneously analyzing and isolating the components of virions.

Automated N-Terminal (Edman Degradation) Protein Sequencing

Long before DNA could be sequenced [21], proteins were being routinely sequenced by using chemicals or enzymes. The original Sanger reagent sequencing, which chemically labeled the amino-terminal residue from a polypeptide which was then partially hydrolyzed and the fragments separated and completely hydrolyzed

(1951), was soon supplanted by the Edman degradation technique which immobilized the carboxy-terminus of the protein and then chemically cleaved the amino-terminal residue so that each step revealed a successive amino acid [9, 10, 17]. These techniques required large amounts of protein and were carried out painstakingly by hand. Automating this process, the protein sequenator developed by Edman in 1967 [22] made protein sequencing more rapid but still required large amounts of protein and the length of sequence that could be read was rather limited. Therefore, to sequence proteins, one had to break them up into small polypeptides by cleaving the protein with proteases and then sequencing the peptides. By manually comparing and knitting together the short sequences produced from several protease digests, a sequence of the whole protein could be obtained. Modern protein sequencing machines use an approach pioneered by Hunkapiller and Hood, who developed the protein microsequencer in 1978 [23]. With this new tool, it was possible to sequence 77 amino acids from using 5 mg of antibody light chain. Much longer sequence reads from much less material in an automated format revolutionized protein sequencing, vastly increasing the speed and ease of sequencing.

Today, the chemically based microsequencer is still used. Its disadvantages are that it requires much more protein than ms/ms spectrometry sequencing. Also, it works best on a purified protein; the presence of other proteins causes other amino acids to be present in the reads. While low levels of contaminants can be tolerated, being less abundant than the proteins of interest, not all amino acids yield the same signal; thus, it is possible to have a weakly yielding residue in the protein at the same position as a strong amino acid signal in the contaminant which could lead to a miscall. Thus, protein sequencers cannot analyze complex mixtures of proteins. Also, the chemistry of Edman degradation sequencing requires a free amino group on the end. Therefore, proteins with modified amino-termini, e.g., those with acyl modifications such as HIV-1 Gag, cannot be directly sequenced [23]. To sequence the remainder of the protein, this modified amino acid must be removed, typically by a downstream protease cleavage to produce a “free” amino-terminal end. This is a real shortcoming as roughly half of the proteins in the cell have blocked amino-termini. Furthermore, “reading” the sequence output requires some skill that must be learned from experience: how to adjust the results for differing amino acid yields and looking for potentially modified amino acids, e.g., ubiquitinated proteins or cyclic proteins, which make interpretation difficult. On the positive side in comparison to MS/MS spectrometry sequencing, the length of the sequencing reads is much longer and is direct. The sequence is obtained without any bioinformatical statistical fitting to a protein sequence database that could miss proteins due to allelic polymorphism, mutations, or posttranslational modifications not noted in the database. Also, some proteins are cleaved into too small a polypeptide to be statistically significant enough to be reported. This is especially true with basic proteins and trypsin digestion, a commonly used protease for fragmentation. Finally, the coverage of an MS/MS spectrometry sequence is only as good as its database and its digestion procedure. Despite its low throughput and other shortcomings, the protein microsequencer still remains an important protein analysis tool.

Amino Acid Analysis

Amino acid analysis analyzes the amino acids present in a protein by first acid hydrolyzing the protein and then separating and detecting the individual amino acids. This technique, once done manually with thin layer chromatography, is now carried out in an automated format which greatly increases speed and accuracy. This method provides the most precise way to analyze the amino acid makeup of a protein. Amino acid analysis also is the most accurate method to measure the absolute amount of protein in a sample. However, a downside to this technique is that it requires relative larger amounts of proteins than some other methods and samples that are very pure. In HIV-1 biochemistry, this was used to identify the amino acid composition of the HIV virion proteins before sequencing. With the roster of amino acids found in a protein, the sequencing is more easy to read and assemble since one only needs to arrange the amino acids like a puzzle rather than discover and then order them. Sequence ambiguities such as missing amino acids and miscalls are more easily resolved. Also, once the sequence is known, then modified residues can be identified by their differing elution profile on the HPLC separation. This technique is also useful for quality control of recombinant and purified proteins.

These classical techniques have served us well over the last few decades and are still very relevant to the study of proteins in HIV-1 particles. Looking to the future, there will be a long partnership between the old-school "proteomics before there was proteomics" and proteomics.

Acknowledgments This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services; nor does the mention of trade names, commercial products, or organizations imply endorsement by the US government.

References

1. Baltimore D. RNA-dependent DNA, polymerase in virions of RNA tumour viruses. Nature. 1970;226(5252):1209–11.
2. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of *Rous sarcoma* virus. Nature. 1970;226(5252):1211–3.
3. Vogt P. Historical introduction to the general properties of retroviruses. In: Coffin J, Hughes S, Varmus H, editors. Retroviruses. Plainview, NY: Cold Spring Harbor Laboratory Press; 1997. p. 1–26.
4. Vogt P. Historical introduction to the general properties of retroviruses. Retroviruses 1997. [cited]. http://www.ncbi.nlm.nih.gov/books/NBK19388/
5. Online M-W. 2014. [cited]. http://www.merriam-webster.com/dictionary/retrovirus
6. Crick FH. On protein synthesis. Symp Soc Exp Biol. 1958;12:138–63.
7. Crick F. Central dogma of molecular biology. Nature. 1970;227(5258):561–3.
8. Temin HM. The DNA, provirus hypothesis. Science. 1976;192(4244):1075–80.

9. Sanger F. Chemistry of insulin; determination of the structure of insulin opens the way to greater understanding of life processes. Science. 1959;129(3359):1340–4.
10. Edman P. On the mechanism of the phenyl isothiocyanate degradation of peptides. Acta Chem Scand. 1956;10:761–8.
11. Ott DE. Potential roles of cellular proteins in HIV-1. Rev Med Virol. 2002;12(6):359–74.
12. Ott DE. Cellular proteins detected in HIV-1. Rev Med Virol. 2008;18:159–75.
13. Chertova E, Chertov O, Coren LV, Roser JD, Trubey CM, Bess Jr JW, et al. Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. J Virol. 2006;80(18):9039–52.
14. Coren LV, Shatzer T, Ott DE. CD45 immunoaffinity depletion of vesicles from Jurkat T cells demonstrates that "exosomes" contain CD45: no evidence for a distinct exosome/HIV-1 budding pathway. Retrovirology. 2008;5(1):64.
15. Wikipedia. Isoelectic point. [cited]. http://en.wikipedia.org/wiki/Isoelectric_point
16. Shapiro AL, Vinuela E, Maizel Jr JV. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. Biochem Biophys Res Commun. 1967;28(5):815–20.
17. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970;227(5259):680–5.
18. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis, 1975. Biotechnology. 1992;24:122–39.
19. Soutnern E. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 1975;98:503–71.
20. Scott RP, Kucera P. Use of microbore columns for the separation of substances of biological origin. J Chromatogr. 1979;185:27–41.
21. Donelson JE, Wu R. Nucleotide sequence analysis of deoxyribonucleic acid. VI. Determination of 3′-terminal dinucleotide sequences of several species of duplex deoxyribonucleic acid using *Escherichia coli* deoxyribonucleic acid polymerase I. J Biol Chem. 1972;247(14):4654–60.
22. Edman P, Begg G. A protein sequenator. Eur J Biochem. 1967;1(1):80–91.
23. Hunkapiller MW, Hood LE. Direct microsequence analysis of polypeptides using an improved sequenator, a nonprotein carrier (polybrene), and high pressure liquid chromatography. Biochemistry. 1978;17(11):2124–33.