DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 10(2), 289-304 (1984)

ANALYSIS AND CONTROL OF PROTEIN AND POLYPEPTIDE DRUGS[1]

Robert V. Smith and Makau P. Lee
Drug Dynamics Institute
College of Pharmacy
University of Texas at Austin
Austin, TX 78712-1074

"We are now in a period of especially rapid progress in applied biology. Important useful advances have already occurred employing recombinant DNA and hybridomas. Synthetic human insulin is being sold commercially and other major pharmaceuticals for human and domestic animal care are being tested... Prospects are excellent that viral diseases soon will be conquered by use of interferon or vaccines."

Philip H. Abelson[2]

The use of protein and polypeptide drugs for the prevention and treatment of diseases is not new. Immunological agents in particular have been used medically since 1798 when Edward Jenner developed a smallpox vaccine from cattle infected with cowpox.[3]

0363-9045/84/1002-0289$3.50/0

Up until the 1970's proteinaceous drugs were developed primarily from materials isolated from human or animal tissues, or through relatively classical microbiological processes. In the last 10 years, however, there have been remarkable advances in the development of molecular synthesizers and in recombinant DNA (rDNA) techniques.[4] These methods should foster a renaissance of interest in protein and polypeptide drugs.

Official Proteinaceous Drugs

The United States Pharmacopeia[5] (U.S.P.) lists over 50 proteinaceous drugs and drug products (see Table 1). Many of these drugs are defined as biologicals and require a batch-by-batch certification by the Food and Drug Administration (FDA).[5] The certification process has been in force since the passage of the Public Health Service Act (58 Stat. 682) in 1944 and uniformly involves bioassay procedures.

Official proteinaceous drugs other than those categorized as "biologics" are quantitated by bioassay methods and in some instances, chemical determinations. As indicated in Table 2, the chemical assays tend to be non-specific. Indeed, the sophistication of assays for proteinaceous drugs is currently where assays for low molecular weight drugs were over 10 years ago. At that time, the majority of official monographs (for non-proteinaceous drugs) prescribed titrimetric or other non-specific assays. Analogously, today, no U.S.P. assays for proteinaeous drugs include specific measurements of identity and purity; though, measures of specific activity of drugs with

TABLE 1. Official Proteinaceous Drugs Evaluated by Biological Assays

Antihemophilic Factor[a]	Glucagon	Poliovirus Vaccine Preparations[a]
Anti-Human Globulin Serum[a]	Heparin Preparations[a]	Polymixin B Sulfate
Antirabies Serum[a]	Hepatitis B Immune Globulin[a]	Protamine Sulfate
Antivenin[a]	Immune Globulin[a]	Rho (D) Immune Globulin[a]
Bacitracin Preparations	Influenza Virus Vaccine[a]	Red Blood Cells and Whole Blood[a]
BCG Vaccine[a]	Insulin	Rocky Mountain Spotted Fever Vaccine[a]
Blood Grouping Serums[a]	Iodinated Albumin Products[a]	Rubella Virus Vaccine Live[a]
Botulism Antitoxin[a]	Lymphogranuloma Venereum Antigen	Schick Test Control[a]
Capreomycin Sulfate	Lypressin Nasal Solution	Tetanus Antitoxin and Toxoid[a]
Cholera Vaccine[a]	Measles, Mumps, and Rubella Viruses[a]	Thrombin[a]
Chorionic Gonadotropin	Menotropins	Tuberculin[a]
Coccidioidin[a]	Mumps Skin Test Antigen[a]	Typhoid Vaccine[a]
Corticotropin	Mumps Virus Vaccines[a]	Typhus Vaccine[a]
Diptheria Toxoid[a]	Oxytocin Preparations	Vaccinia Immune Globulin[a]
Factor IX Complex	Platelet Concentrate[a]	Vasopressin Injection

[a] Biologic product requiring lot-per-lot certification by FDA (U.S.P., pp. 993-994).

TABLE 2

Official Proteinaceous Drugs

Evaluated by Chemical Assays

Chymotrypsin[a]
Gelatin Products[b]
Hyaluronidase Injection[c]
Insulin[d]
Iodinated Albumin Products[e]
Pancreatin and Pancrelipase[f]
Thyroglobulin and Thyroid[g]

[a] Enzymatic assay; UV spectrophotometric measurement step.

[b] Identified by colorimetric and turbidimetric tests; evaluated for microbial contamination and tested for: residue on ignition, odor and water-insoluble substances, sulfur dioxide, arsenic, and heavy metals.

[c] Colorimetric assay of tyrosine content following hydrolysis.

[d] Product assayed for nitrogen content by Kjeldahl determination; tested for residue on ignition and zinc content where relevant.

[e] Assayed for radiochemical purity by paper chromatography and radiochromatographic scanning.

[f] Colorimetric assays for amylase and protease activities; pH-stat procedure for lipase activity.

[g] Iodometric titration.

enzymatic activity are included where appropriate (see Table 2). An assay for natural bovine and pork insulins based on high performance liquid chromatography (HPLC) has, however, been proposed in the Pharmacopeial Forum.[6] For other official drugs, and for products that will be developed during the remainder of the 20th century, specific methods will become desirable.

Drugs Derived Through Recombinant DNA Research

Initial research efforts with the production of proteins using recombinant DNA (rDNA) techniques occurred about 10 years ago.[7,8] The short history of this research is marked by controversy over the potential dangers of handling common gut bacteria (e.g., Escherichia coli) containing foreign genetic material.[9] Recent safety appraisals[10] and Federal Guidelines[11] for conducting rDNA research provide more realistic evaluations of apparent dangers. The moderated safeguards assure greater activity in the rDNA field.

In 1982, human insulin prepared by rDNA-technology became the first drug of its type to be approved for use by the drug regulatory agencies in the United Kingdom, the Netherlands, West Germany and the United States.[10] The recombinant insulin is identical to that produced in humans as determined by HPLC, gel electrophoresis, isoelectric focusing, circular dichroic (CD) spectrophotometry, amino acid analyses, radioreceptor assay, radioimmunoassay (RIA) and rabbit hypoglycemic assay.[12] Furthermore, the purified product appears to have negligible contamination by foreign proteins (e.g., E. coli endotoxin) as determined by HPLC and the Limulus amebocyte lysate assay.[12]

TABLE 3

Proteinaceous Drug Candidates for Production By rDNA Methods[a]

Drug	Number of Amino Acid Residues	Approximate Molecular Weight $x10^3$
Active Fragment of Corticotropin (ACTH)[b]	24	3
Big Gastrin	34	4
Calcitonin[b]	32	3.4
Corticotropin	39	4.5
Endorphins	31	3.5
Gastric Inhibitory Polypeptide	43	5
Glucagon[b]	29	3
Growth hormone[b]	191	22
Interferons	167	20[c]
Nerve Growth Factor	118	13
Parathyroid Hormone	84	9.5
Placental Lactogen	192	22
Proinsulin	82	9
Prolactin	198	23
Secretin[b]	27	3
Thymopoietin	49	5.5
Thymosin-α_1	28	3.1
Vasoactive Intestinal Peptide	28	3

[a] Some of the drugs listed (e.g., certain interferons, human growth hormone and proinsulin) have already been synthesized by rDNA methods. None have been approved for marketing in the United States.

[b] Currently used in medical practice but obtained from natural sources or through automated protein synthesizers which can be used to produce economically polypeptides with amino acid residues up to about 32.[5]

[c] Non-glycosylated interferon-α.

The success in producing human insulin by rDNA methods is a portent of research that will lead to the preparation and testing of greater numbers of biologically active polypeptides and proteins during the next 20 years. Some likely candidates[4] are indicated in Table 3.

Analytical Chemical Challenges and Approaches

Qualitative and quantitative methods of analysis are required to control rDNA-produced proteinaceous drugs. The procedures should permit identification of the principal constituent(s) and should help determine if miscoding has occurred which can lead to proteins with altered amino acid sequences. Qualitative methods should also allow detection of trace impurities of the type listed in Table 4. The sources of the potential trace impurities are also given in Table 4 and reflect the complex nature of drug production by rDNA methods.

The potential trace impurity problem should be put into perspective. Numerous biological products (e.g., natural porcine insulin, vaccines) containing some foreign proteins have been used clinically for years without serious adverse effects. Recent experiences with human insulin produced by rDNA methods suggest that foreign protein contamination can be minimized (i.e., $\bar{<}$ 3 percent).[12] Contamination of drug products by low molecular weight impurities has been prevented for years by major pharmaceutical companies.

Analytical methods should permit quantitation of the rDNA-derived proteinaceous constituent and provide measures of

TABLE 4

Potential Trace Impurities in rDNA-derived Drugs

Impurity	Source
Nucleic acids (DNA and RNA) and microbial proteins including proteases	Rupture of drug-producing microorganism during isolation and purification procedures
Mouse IgG	Used during production of monoclonal antibodies
Monoclonal antibodies	Used in immunoaffinity purification steps
Antibiotics	Added to culture media to select for growth of rDNA plasmid-containing organisms.
Surfactants, solvents	Used in isolating and purifying rDNA-derived drug product.
Protease inhibitors and/or hydrolysis products	Prevent hydrolysis of proteinaceous drug during isolation.

homogeneity. Most importantly, the homogeneous product must be controlled for biological activity and this will require biological or hybrid biological-chemical methods. Developed procedures should possess accuracy and precision approaching that expected by industrial scientists and FDA officials for pharmaceutical products.[13]

A traditional approach to the analysis of proteinaceous drugs and their trace impurities might involve use of a series of specific tests. Thus, protein content could be assayed by the Lowry method [14] or newer procedures using Coomassie Brilliant Blue.[15,16] Trace determinations of nucleic acids collectively (i.e., DNA and RNA) and DNA specifically could be accomplished through fluorometric procedures using ethidium bromide[17] and 4',6-diamino-2-phenylindole[18,19] or a related benzimidazole[20], respectively. Determination of trace low molecular weight impurities could be affected by published gas chromatographic (GC)[21] and/or high performance liquid chromatographic (HPLC) methods.[22] The series of chemical procedures would be "capped off" by a bioassay step which would quantify the active portion of the total proteinaceous content. This "train" of methods would obviously be time consuming and expensive to perform. Without modern chromatographic methods, however, there would be no alternative. We believe that recent advances in HPLC, isoelectric focusing and gel electrophoresis provide more efficient approaches to the analysis and control of rDNA-derived proteinaceous drugs.

Initial qualitative screening could be affected through gel

electrophoresis[23-25] or isoelectric focusing, or a combination of the two as used in analytical studies with rDNA-derived human insulin.[12] Differential staining methods are available[26] for detecting the various biopolymeric materials which may occur in proteinaceous drug samples or drug dosage forms (see Table 3). Proteinaeous components can also be sensitively detected (~ 1 ng) with silver stains.[27]

Other qualitative procedures for rDNA-derived proteins include protease mapping, amino acid analyses, very high resolution nuclear magnetic resonance (NMR) spectrometry, and circular dichroic (CD) spectrophotometry. Protease mapping involves controlled hydrolysis with a protease such as trypsin[28] (i.e., resulting in so-called tryptic mapping) or other more selective proteases and resolution of resulting peptide fragments by HPLC or comparable technique. The HP-liquid chromatograms produced with proteins of moderate molecular weight (e.g., 20,000 Daltons) are complex and reproducible hydrolysis conditions can be difficult to establish.

Amino acid analysis involving complete hydrolysis and analysis of all component amino acids is readily performed with automated equipment but is of limited specificity with respect to the intact protein. Amino acid sequencing involving one-by-one removal and analysis of component amino acids[29,30] can provide important information on primary structure and would detect miscoding. This technique, however, would not be cost effective for routine control procedures.

Very high resolution NMR methods are potentially a means of

determining primary, secondary and tertiary structures of rDNA proteins.[31] At present, however, necessary equipment ($\bar{>}$ 400 MHz) for moderate to high molecular weight proteins is not readily available, is very expensive, and is untested on a wide array of proteinaceous drugs. The tertiary structure of polypeptides and proteins can also be determined by CD spectrophotometry as was used with human insulin.[10,12]

HPLC offers methods of choice for the identification and quantitation of rDNA-derived proteins, degradation products and possibly denatured molecules. Recently developed ion exchange (IE), reverse phase (RP) and size exclusion (SE) HPLC column materials and separation routines have great potential usefulness in protein analyses.[32-35] The resolution of insulins from various animal sources by RP-HPLC exemplifies the powerfulness of the technique. Chance et al.[12] separated rabbit, human and porcine insulins on an octadecylsilyl-bonded phase column using a ternary isocratic solvent system composed of 2-methoxyethanol-acetonitrile-phosphate buffer combinations. The three insulins have molecular weights around 5,700 Daltons and differ in structure by one terminal amino acid (i.e., the terminal carboxy- amino acids of the A-chains of rabbit, human and pork insulins are serine, threonine and alanine, respectively). Even more remarkably, RP-HPLC was used to develop a separation of beef and sheep insulins which differ by addition of a single glycine residue in the A-chain (i.e., 9-position) of the hormone from sheep. The insulins are notably lower in molecular weight than

many other proteinaceous drugs of interest. The impressive selectivity achieved with RP-HPLC however, suggests that analogous separations can be affected with higher molecular weight proteins. Ultimately this could led to methods that detect miscoding.

Available detection systems for HPLC (e.g., UV-visible, fluorescence) are readily adaptable to the measurement of proteins. In instances where higher sensitivities are required, post column reactions involving fluorophore-generating reagents (e.g., fluorescamine) can be used.[36]

Surface effects or organic solvent modifiers may denature proteinaceous drugs during HPLC.[34,35] If precautions are taken, however, denaturation can be minimized. This allows coupling of the HPLC separation and chemical detection steps with a biological measurement step. The latter may be accomplished by collection of chromatographic peaks followed by immunoassay or comparable bioassay procedures. Alternatively, column effluents could be mixed in post column reactors with monoclonal antibodies or other complexing agents for subsequent detection by nephelometry. Another coupling procedure might involve routing eluates through secondary immunoaffinity columns. Retention on the latter column might provide a measure of biological potency. These alternative methods have not been tried but represent possible approaches to efficient chemical-biological analyses.

Methods developed along the lines outlined above will require careful evaluation. Of particular importance will be correlations of results with standard bioassay procedures.

Summary and Conclusions

The problem of analysis and control of proteinaceous drugs has been reviewed. Unique difficulties associated with drugs derived from rDNA technology have been highlighted and approaches for chemical and biological assays proposed. Eventually, readily adoptable chemical methods will be developed to control the primary, secondary and tertiary structures of biologically active proteins. These methods could eliminate the need for Federal certification of many "biologics" much as similar requirements for antibiotics were abandoned after the development of specific analytical chemical procedures.

ACKNOWLEDGEMENTS

The authors are grateful to Drs. Jane Sheridan and Frances Bogdansky for their encouragement and support. Financial support from Hoffmann-LaRoche Inc. is also gratefully acknowledged.

REFERENCES

1. Presented by Robert V. Smith at the 22nd Annual International Industrial Pharmacy Conference, Austin, Texas, February, 1983. Peptides are defined as amino acid polymers with molecular weights (MWs) less than 1,000; polypeptides have MWs from 1,000 to 5,000 and proteins have MWs > 5,000. We use the word proteinaceous to denote polypeptides and proteins collectively.
2. P. Abelson, Science, 219, 609 (1983).
3. G. Migaki, Lab. Animal, 10 (2), 30 (1981).
4. "Impact of Applied Genetics -- Microorganisms, Plants, and Animals," Office of Technology Assessment, Congress of the

United States, OTA-HR-132, U.S. Printing Office, Washington, D.C., 1981, pp. 57-82.

5. The United States Pharmacopeia, 20th Revision, United States Pharmacopeial Convention, Rockville, MD, 1980.
6. Pharmacopeial Forum, 2595, Jan.- Feb. (1983).
7. D. A. Jackson, R. H. Symons, and P. Berg, Proc. Natl. Acad. Sci. U.S.A., 69, 2904 (1972).
8. S. N. Cohen, A.C.Y. Chang, H. W. Boyer, and R. B. Helling, Proc. Natl. Acad. Sci. U.S.A., 69, 3240 (1973).
9. D. S. Fredrickson, in "Recombinant DNA and Genetic Experimentation," J. Morgan and W. J. Whelan, Eds., Pergamon Press, New York, 1981, pp. 151-156.
10. I. S. Johnson, Science, 219, 632 (1983).
11. Guidelines for Research Involving Recombinant DNA Molecules. Fed. Register, 47 (167), 3804, August 27, 1982.
12. R. E. Chance, E. P. Kroeff, J. A. Hoffmann, and B. H. Frank, Diabetes Care, 4, 147 (1981).
13. W. Horwitz, Anal. Chem., 54, 67A (1982).
14. O. H. Lowry, N. H. Rosebrough, A. L. Farr, and R. J. Randall, J. Biol. Chem., 193, 265 (1951).
15. M. M. Bradford, Anal. Biochem., 72, 248 (1976).
16. T. Spector, Anal. Biochem., 86, 142 (1978).
17. J. B. LePecq and C. Paoletti, Anal. Biochem., 17, 100 (1966).
18. W. C. Russel, C. Newman, and D. H. Williamson, Nature, 253, 461 (1975).
19. J. Kapuscinski and B. Skoczylas, Anal. Biochem., 83, 252 (1977).

20. C. F. Brunk, K. C. Jones, and T. W. James, Anal. Biochem., 92, 497 (1979).

21. B. J. Kline and W. H. Soine, in "Pharmaceutical Analysis -- Modern Methods," Part A., J. W. Munson, Ed., Marcel Dekker, New York, 1981, pp. 1-103.

22. L. R. Snyder and J. J. Kirkland, "Introduction to Modern Liquid Chromatography," 2nd ed., Wiley, New York, 1979.

23. P. H. O'Farrell, J. Biol. Chem. 250, 4007 (1975).

24. E. M. Southern, J. Mol. Biol., 98, 503 (1975).

25. H. Towbin, T. Staehelin, and J. Gordon, Proc. Natl. Acad. Sci. U.S.A., 76, 4350 (1979).

26. M. R. Green, J. V. Pastewka, and A. C. Peacock, Anal. Biochem., 56, 43 (1973).

27. S. J. Tarnowski, Pharm. Tech., 7 (2), 70 (1983).

28. R. Wetzel, L. J. Perry, D. A. Estell, N. Lin, H. L. Levine, B. Slinker, F. Fields, M. J. Ross, and J. Shively, J. Interferon Res., 1, 381 (1981).

29. D. G. Klapper, C. E. Wilde III, and J. D. Capra, Anal. Biochem., 104, 464 (1980).

30. R. Somack, Anal. Biochem., 104, 464 (1980).

31. G. C. K. Roberts and O. Jardetzky, Advan. Protein Chem., 24, 448 (1970).

32. M. T. W. Hearn, Adv. Chromatogr., 20, 1 (1982).

33. R. V. Lewis and D. DeWald, J. Liq. Chromatogr., 5, 1367 (1982).

34. F. E. Regnier and K. M. Gooding, Anal. Biochem., 103, 1 (1980).

35. T. W. Hearn, F. E. Regnier, and C. T. Wehr, Amer. Lab., 14 (10), 18 (1982).

36. S. Stein and J. Moschera, in "Methods in Enzymology," Vol. 79, S. Pestka, Ed., Academic Press, New York, 1981, pp. 7-16.