DEPARTMENT OF HEALTH & HUMAN SERVICES Public Hedh Service

P011

Food and Drug Administration
Bethesda MD 20892

Date April 6, 1992

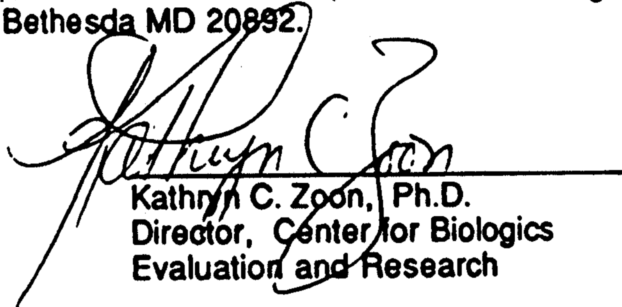
From Center for Biologics Evaluation and Research
Food and Drug Administration

Subject Supplement to the Points to Consider in the Production and Testing of New Drugs and Biologic& Produced by Recombinant DNA Technology: Nucleic Acid Characterization and Genetic Stability

To Manufacturers of Biological Products and Other Related Parties

This Supplement has been developed for recombinant DNA protein products and is not a regulation nor is it a guideline, but represents the current thinking of the Center for Biologics Evaluation and Research (CBER) staff.

We intend to continually revise and update this document in order to improve its usefulness, effectiveness, and practicality. We invite your review and comments on this document. Your comments should be addressed to the Director, Center for Biologics Evaluation and Research, Food and Drug Administration, 6600 Rockville Pike, Bethesda MD 20892.



Kathryn C. Zoon, Ph.D.
Director, Center for Biologics
Evaluation and Research

I. Introduction

This document presents guidance regarding the characterization of the expression construct for the production of recombinant DNA protein products in eukaryotic and prokaryotic cells. The expression construct is defined as the expression vector containing the coding sequence of the recombinant protein. Characterization Of the expression construct was one of the issues addressed in the "Points to Consider in the Production and Testing of New Drugs and Biologicals Produced by Recombinant DNA Technology", 1985. Since that time there have been advances in analytical technologies for protein and nucleic acid characterization, which have occurred in parallel with advances in fermentation and cell culture technologies. This document is intended to extend and clarify the types of information that are considered valuable in assessing the structure of the expression construct used to produce recombinant DNA proteins. In developing this document, we have attempted to take into consideration recent advances in biochemical technology, past history with different production systems, information derived from international meetings, and comments submitted to the FDA. In particular, these issues were discussed in detail at the "1991 International Workshop on Continuous Cell Lines: Current Issues'. Scientists at the Center for Biologics Evaluation and Research (CBER) believe that the supplemental guidance in this document reflects commonly held positions within the scientific community and that this document clarifies the testing to be performed to characterize nucleic acids and assess genetic stability. CBER also recognizes that this area has been one of special concern to the industry and invites comments on this document that can be considered in subsequent revisions to the Points to Consider in the

II. Rationale for Nucleic Acid Testing of the Expression Construct

Recombinant proteins are produced in living cells which can undergo mutations that could alter the properties of a protein with potential adverse consequences to patients. No single experimental approach can be expected to detect all possible modifications to a protein. Protein analytical techniques are useful in assessing -many structural features of the expressed protein as well as changes due to post-translational modification systems such as those affecting protein proteolytic processing, glycosylation, phosphorylation, and **acetylation** and other alterations to the protein, e.g., deamidation, oxidation. However, analysis of the final purified protein may not detect every change in a protein's structure that could result from a mutation in the protein **coding** sequence. For example, conservative amino acid substitutions may be missed by **peptide** mapping techniques and yet could cause structural changes in the protein. Some changes might not affect **in-vitro** activity but could lead to significant **in-vivo** effects, such as alterations in pharmacodynamics, pharmacokinetics, or induction of antibody responses. In a population of cells making large amounts of a foreign protein, mutant cells making an altered protein or reduced amounts of protein may have *a* growth advantage and thus increase their representation in the population after multiple generations. Evaluation of the final purified protein may miss variants that are synthesized but normally removed by downstream processing. A breakdown in the efficiency of purification steps might allow such variants to contaminate the final

purified protein. Mutations in the gene(s) encoding the protein, however, could theoretically be detected at the DNA or RNA level. For the above reasons, a combination Of methods, which includes analysis of both the purified protein and nucleic acid, should be chosen carefully in order to ensure product consistency.

Data derived from analysis of the purified protein utilizing analytical biochemical techniques are relatively straightforward to interpret and have been discussed in the "Points to Consider in the Production and Testing of New Drugs and Biologicals Produced by Recombinant DNA Technology" 1985. In some cases nucleic acid test results are more difficult to interpret with respect to the final product; however, such data can provide valuable information that is complementary to data derived from analysis for the purified protein. Nucleic acid analysis can be used to verify the coding sequence of the protein and the physical state of the expression construct. Verification of the coding sequence is not intended to detect low levels of variant sequences, but rather to validate that the predominant protein species prior to purification will have the correct amino acid sequence following the manipulations of cells during full scale production. Where the production cells have multiple integrated copies Of the expression construct, not all of which may be transcriptionally active, examination of the transcription product itself by mRNA or cDNA analysis may be more appropriate than genomic DNA analysis. Analytical approaches that examine a bulk population of nucleic acid, such as those performed on material amplified by the polymerase chain reaction, may be preferable to approaches depending on selection of individual DNA clones.

The following sections describe information that should be supplied regarding the characterization of the expression construct during the

development and validation of the cell expression clone for production. These recommendations are based on current technology, and it is anticipated that they will be modified as newer techniques are developed and as our knowledge grows in this area.

III. Characterization of Expression Constructs

A. <u>Initial Cell Clone</u>

<u>DNA Clone</u> - The manufacturer should describe the method used to prepare the DNA coding for the protein, including the cell and origin of the source nucleic acid.

The steps in the assembly of the expression construct should be described in detail. This description should include the source and function of the component parts of the expression construct, e.g., origins of replication, antibiotic resistance genes, promoters, enhancers, whether or not the protein is being synthesized as a fusion protein. Restriction endonuclease digestion maps illustrating the sites used in the construction of the expression construct and sites used in identification of DNA fragments should be provided.

A complete nucleotide sequence analysis of the coding region for the protein in the expression construct should be performed. The sequence analysis should be provided and include a complete annotation designating all of the important sequence features, e.g., sites involved in the construction, alterations from the original sequence. If the protein coding region contains more information than coding sequence such as introns or flanking sequences, these additional sequences should be characterized.

Cell Expression Clone • The host cell system that will generate the protein is designed to be compatible with the expression construct. It is, therefore, important that a description of the source, phenotype, and genotype of the host cell be provided, including literature references. The host cell should be thoroughly characterized as described in the "Points to Consider in the Characterization of Cell Lines to Produce Biological Products", 1987. A description of the method of transfer of the expression construct into the host cell should be provided. In addition, methods and criteria used to amplify the expression construct and select the cell clone for production should be completely described. The copy number and physical state of the expression construct also should be determined.

B. Master

The Master Cell Bank (MCB) is a collection of cells of uniform composition stored in aliquots under defined conditions, from which all subsequent cell banks are made as discussed in the "Points to Consider in the Characterization of Cell Lines to Produce Biological Products', 1987. In most cases, a single host cell containing the expression construct is cloned to give rise to the MCB. The cell cloning history, including relevant dates of creation and methodology, should be described. If a new MCB is to be generated by expression construct transfer into host cells, followed by clonal selection, then acceptance criteria for both the new clone and the protein produced by the clone should be described. The identity and purity of the cells in each MCB should be assured by testing appropriate phenotypic markers, e.g., auxotrophy and antibiotic resistance. Each MCB should be characterized for adventitious agents, if appropriate, as described in the "Points to Consider in the

Characterization of Cell Lines to Produce Biological Products', 1987. The MCB should be tested periodically to ensure the viability of the cells.

The integrity of the expression construct in the MCB should be determined at least once for each MCB. The nucleotide sequence encoding the protein should be verified. For cells with a single genomic copy of the gene coding for the protein, sequencing of total genomic DNA may be preferable to sequencing a single genomic clone. For cells with multiple chromosomal insertions of the expression construct, the nucleotide sequence encoding the product should be verified by direct sequencing of the mRNA or cDNA. For extrachromosomal expression systems the expression construct should be isolated and the nucleotide sequence encoding the product should be verified without further cloning. The expression construct in the MCB should be analyzed by restriction endonuclease mapping for copy number, for large insertions and deletions, and for the number of integration sites (for cells with multiple chromosomal integrations). For extrachromosomal expression systems, the percent of host cells retaining the expression construct should be determined under selected and non-selected growth conditions. The above testing should also apply to the establishment of new MCBs.

c. Manufacturers Cell Bank (MWCB)

The Manufacturers Working Cell Bank (MWCB) is derived by expansion of one or more ampules of the MCB as described in the "Points to Consider in the Characterization of Cell Lines to Produce Biological Products", 1987. Each MWCB is of uniform composition and is stored in aliquots under defined conditions. The production of the MWCB should be described in detail

including methods and reagents used during culture, number of cell doublings from the MCB, and storage conditions. At time of storage the identity of the MWCB should be assessed by restriction endonuclease mapping of the expression construct for copy number and for insertions or deletions. In addition, where appropriate, the MWCB should be identified by phenotypic characterization, e.g., auxotrophy, antibiotic resistance.

D. End of Production Cells

End of Production cells are derived by expansion from the MWCB and are acquired at the end of an actual production run using full scale manufacturing conditions. The integrity of the expression construct in End of Production cells should be determined once for each MCB as described in Section B, Master Cell Bank. Tests carried out on End of Production cells from pilot scale facilities should be repeated on End of Production cells from full scale manufacturing facilities.

It is important to note that tests on End of Production cells described above are not intended to be carried out at the end of each production run. At the end of each production cycle, cells should be analyzed for relevant phenotypic or genotypic markers, and for contamination by adventitious agents as described in the "Points to Consider in the Characterization of Cell Lines to Produce Biological Products", 1987. Data on the consistency of yield from full-scale culture should be maintained, and criteria for the rejection of culture lots should be established.

IV. Conclusion

The characterization of the expression construct and the purified protein is important to ensure the consistent production of a recombinant **DNA** derived product. At present, it is believed that analytical data derived from either nucleic acid testing or protein structural testing **<u>alone</u>** do not allow for a complete evaluation of the identity and purity of a recombinant protein product. As technology advances, CBER will reassess this position and revise this document.