



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service  
National Institutes of Health

September 14, 1998

SOLICITATION  
SEP 15 1998

Office of Technology Transfer  
National Institutes of Health  
6011 Executive Boulevard  
Rockville, MD 20852

Box 8  
Commissioner of Patents and Trademarks  
Patent and Trademark Office  
Washington D.C. 20231

U.S. PATENT & TRADEMARK OFFICE

Attention: Scott A. Chambers, Associate Solicitor

**Comments on Interim Guidelines on the Written Description Requirement**

The written comments presented herein represent the views of the National Institutes of Health (NIH). The NIH is the lead agency within the Public Health Service (PHS) in matters of technology transfer. In addition to providing patent and licensing services to all Institutes and Centers comprising the NIH, PHS lead agency status further encompasses coordinating and facilitating technology transfer policy functions with the Centers for Disease Control and Prevention (CDC) and the Food and Drug Administration (FDA). Central responsibility within NIH for these technology transfer functions has been delegated to the Office of Technology Transfer (OTT).

**Summary**

The U.S. Patent and Trademark Office (PTO) has set forth a workable outline for analyzing applications for compliance with the written description requirements. There appear to be significant deficiencies, however, in analyzing the proper relationship of the preamble, transition phrase, and claim body to determinations of genus versus species claims. This determination is particularly critical to the application of the open-ended transition phrase "comprising" to claims involving nucleic acids and amino acids. A number of examples are presented to clarify this relationship. Due to the highly controversial nature of Expressed Sequence Tag (EST) applications related to its potentially devastating impact on the biotechnology community and the relevancy of these written description considerations to the patentability of EST claims employing "comprising" language, the NIH requests that this issue be addressed specifically in the final guidelines.

**Introduction and Background to Federal Transfer of Biotechnology:**

**A) Legislative Mandate for Federal Technology Transfer**



The Bayh-Dole Act of 1980, Pub. L. No. 96-517, 94 Stat. 3015, as amended, permits recipients of federal grants and contracts to retain intellectual property title to their inventions. This act also permits exclusive licensing of Government-owned inventions. In October 1986, Congress enacted the Federal Technology Transfer Act (FTTA), Pub. L. 99-502, 100 Stat. 1785, which amended the Stevenson-Wydler Innovation Act of 1980. The FTTA, as amended, stimulates transfer of Government-owned technology by offering incentives to both federal laboratories/scientists and collaborating partners in universities, foundations (both profit and non-profit), or private industry. With regard to intramural research, the FTTA obliges government scientists to report inventions having commercial or health benefit potential for transfer to the private sector. To facilitate this obligation, the FTTA provides incentives comprising cash awards and distribution of a portion of licensing royalties back to the laboratory and inventors.

#### **B) NIH Advancement of the Technology Transfer Mandate**

The NIH has engaged in considerable technology transfer activity consequent to the initiatives promulgated by the FTTA. Since fiscal year 1987, the NIH has received over 900 issued patents, executed over 1,300 license agreements, generated about 200 million dollars in royalties, and entered into about 400 Cooperative Research and Development Awards (CRADAs). While significant, these activities reflect the transfer of only a fraction of the cutting-edge invention portfolio generated by the world's preeminent public entity dedicated to the advancement of health care.

Beyond this intramural research contribution, the NIH funds biomedical research at universities and contractor-operated research facilities via research grants and contracts. Funding of extramural grants and contracts constitutes approximately 85% of the 13-plus billion dollar annual budget provided NIH for health research and development. As a result of these two contributory streams, the NIH is the world's leading source and underwriter of biomedical inventions.

A significant proportion of the NIH's intramural research and extramural funding is directed to genomics. This involvement extends to numerous aspects of genomic diagnostics, therapeutics, and sequencing. Consequently, the NIH is a major stakeholder in the genomic arena, and the NIH has commensurate interest in any proposed guidelines related to the examination and patentability of biomedical inventions describing nucleic acid and amino acid sequences.

#### **C) NIH Technology Transfer Policy Issues**

NIH technology transfer policies related to both intramural inventions and funding of extramural research are guided by the NIH mission to advance the public health. When significant intellectual property issues arise within the biotechnology community that

impinge upon that public health mission. NIH exercises its leadership and stewardship role.

A recent intellectual property concern in this regard relates to access by the non-profit research community to research tools. NIH initiated a number of actions toward ameliorating this concern. First, NIH modified its intramural patent and licensing policies to insure that NIH's own technology transfer processes facilitate unencumbered access to research tools. Second, the NIH Director convened a Research Tools Working Group of technology transfer representatives from government, academia, and industry to survey and analyze the issue, and to recommend steps to facilitate the unencumbered flow of research tools and biological materials to, from, and within the research community. Release of the NIH guidelines implementing the recommendations of this work group is scheduled for the end of this calendar year. Third, NIH and DuPont Pharmaceuticals Company recently negotiated a Memorandum of Understanding providing non-profit researchers free access and elimination of "reach through" options in non-commercial research licenses related to a broad based research tool (Cre-lox technology). The agreement satisfies industry's intellectual property and commercial interests; yet satisfies the NIH and the academic community's concerns regarding encumbrance of research tools to the non-profit sector. It is expected that the general terms of this agreement will become a model in the academic and government research communities. Fourth, NIH and various members of the biotechnology and pharmaceutical communities have been engaged in dialogue to find ways to augment NIH's massive genome sequencing initiatives. These initiatives include sequencing and placing into the public domain the entire human genome, as well as libraries of expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs). The NIH considers ESTs and SNPs examples of genomic research tools which need to be made available for unencumbered research to advance the public health.

#### **Summary of Correspondence with the PTO Regarding Concerns Related to Patenting of EST Sequences**

As indicated above, the NIH has public health policy interests in partial DNA sequences (i.e., ESTs) whose primary utility in the research community is as a tool to probe for unknown genes. The NIH has voiced its intellectual property concerns regarding ESTs in various fora, including communications to the U.S. Patent Office (PTO) and the Council of the National Academy of Sciences (NAS). The following is a synopsis of relevant considerations derived from such communications.

Soon after its February 14, 1997 public announcement that the PTO considered ESTs patentable subject matter based upon their utility as probes, the Director of NIH (Dr. Harold Varmus) communicated his deep public health concern that such patents may have a chilling effect within the genomics industry. A health care issue may arise if industry delays or refrains from investing in genomic research and development due to uncertainty surrounding the scope of millions of secret EST claims at the PTO. Beyond

the obvious financial concern of significant research investment potentially being dominated by a substantially *de minimis* technical exercise, many in the industry are fearful of the tangled web of overlapping intellectual properties that will stack against the development of important genomic applications. Dr. Varmus' communication was supplemented by a letter from the NIH Office of Technology Transfer detailing the NIH position on the utility issue of ESTs disclosed as probes for unknown genes. This supplemental letter also discussed enablement (undue breadth) issues raised by potential EST claims containing open-ended "comprising" language which broadens scope by introducing random sequences of indeterminate length. This undue breadth scenario was compared to Examples A and B in the PTO Guidelines and Training Materials regarding enablement in chemical and biotechnical applications. Copies of these letters are enclosed.

On April 2, 1997, Commissioner Lehman responded to these NIH communications. The Commissioner acknowledged the NIH concerns and indicated, "[m]ere allegation of the utility of an EST as a probe without further disclosure is not sufficient to meet the utility and enablement criteria." Commissioner Lehman elaborated that potential EST utilities related to forensic identification, tissue type or origin identification, chromosome mapping, chromosome identification, and tagging a gene of known and useful function. These utilities were indicated to be potentially enabled "if supported by a sufficient disclosure." Related to the scope of EST claims, the Commissioner stated, "[u]nder appropriate and limited circumstances, claims of a perceived broad scope that are adequately supported by the disclosure under 35 USC 112 and the state of the art may be patentable,..."

The above exchange of communications and other issues related to patenting of research tools were discussed at the Council of the NAS. The NAS is a society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Under its charter granted by Congress in 1863, the NAS is mandated to advise the federal government on scientific and technical matters.

The NAS has a long standing interest in the intellectual property aspects of research tools used in molecular biology. Since 1993, the NAS has conducted two major workshops on this issue, including one on ESTs. Further, the NAS has published a 1997 National Research Council report on the subject of research tools.

Pursuant to these discussions, Dr. Bruce Alberts, President of the NAS, also communicated with Commissioner Lehman. Dr. Alberts' June 19, 1997 correspondence reiterated the concerns of NIH, and sought clarification of the Commissioner's statement above regarding the possibility of EST claims of broad scope. Furthermore, Dr. Alberts urged the PTO to question the potential enabled utilities proposed in the Lehman letter to Harold Varmus. Communicating on behalf of the Council of the NAS, Dr. Alberts stated the following:

[d]isclosure of DNA sequence alone is plainly insufficient to enable scientists to use an EST for any of these purposes. Data about the exact chromosomal site from which a DNA fragment arose are needed for mapping; data about unique expression in a particular tissue or physiological state are needed for tissue typing or diagnosis; and data about polymorphism among individuals are needed for forensic uses.

In the PTO response to this NAS representation of the state of art related to the enablement of the indicated EST utilities, Commissioner Lehman indicated the following:

The NAS has urged the USPTO to question whether the EST patent applications have applied a sufficient enabling disclosure regarding exact chromosomal sites, unique expression in a particular tissue, or polymorphism among individuals to enable the use of these DNA sequences for mapping, tissue typing, or forensic use. Considerations such as these are clearly within the scope of 35 U.S.C. §112 and are fully considered in accordance with the *In re Wand's* decision in the enablement determination of every claimed invention.

In each of the above communications from the Commissioner, the PTO appears to acknowledge the relevance of the NIH and NAS legal and scientific positions regarding the utility and potential scope of EST claims. As appropriate, the PTO responses imply the issues would be examined on a case by case basis consistent with the relevant case law and published PTO guidelines on utility and enablement. More recently, however, PTO presentations at various public meetings, such as the 1998 Annual Meeting of the Biotechnology Industry Organization (BIO), indicate a perceptible hardening regarding the EST issue. Despite its predisposition against *per se* rules in the examination process, the PTO appears to be contemplating accepting broad disclosure of any or all of the above identified potential utilities as satisfying the 35 USC 101 requirement for all claimed ESTs. Furthermore, such presentations indicate generalized willingness to apply broad scope "comprising" language in EST claims. NIH finds most disturbing these representations of a potentially evolving policy toward accepting utility and broad claim scope *per se* for EST patent applications despite significant NIH/NAS legal and scientific arguments that should militate against such general considerations and conclusions. The NIH believes developments in case law on written description also militate against the issuance of broad EST claims containing open "comprising" transition language.

It was anticipated that the pending interim guidelines on the written description requirement of 35 USC 112 might shed light on the rationale underlying the PTO's intentions regarding these controversial issues. However, specific mention of this class of invention involving nucleic acid sequences is conspicuously absent from these interim guidelines. The failure to address this subject is particularly disturbing considering the

huge number of ESTs pending at the PTO, and the serious concerns raised about ESTs by varied groups interested in the well being and continued development of the biotechnology community. In addition to the public health issues raised by the NIH and the technical and science policy issues raised by the NAS, BIO, and numerous other biotechnology and pharmaceutical companies have also raised concerns about issuance of broad EST patents

The NIH requests that written description issues related to EST claims be formally addressed, including examples, in the Final Guidelines on Written Description. To the extent there are significantly divergent opinions expressed to the PTO regarding written description issues related to ESTs, it would be appropriate to enumerate these views, as well as the PTO's evaluation of the same in arriving at its final guideline determinations. Toward that end, the NIH submits the following comments related to the written description guidelines generally, as well as their application specifically to EST claims.

### Specific Comments on the Interim Guidelines

#### **A) General Outline of Criteria to be Analyzed**

The PTO is commended for its clarity regarding the basic outline of steps and points for consideration in determining whether a disclosure complies with the written description requirement of Section 112, first paragraph. As indicated, the written description requirement is satisfied when the specification describes the claimed invention in sufficient detail to conclude the inventor had possession of the claimed invention at the time of filing.

The interim guidelines succinctly indicate that a proper analysis requires evaluation of the entire application including the specification and the scope of each claim. This evaluation is conducted from the perspective of one skilled in the art at the time the application was filed. Each claim is given its broadest reasonable interpretation, and all parts of the claim (i.e., preamble, transitional phrase, and body) are considered. Also analyzed are the field of the invention and the level of predictability in the art; where the level of predictability in the art is inversely related to the amount of disclosure necessary to demonstrate possession of the claimed invention. It is noted that this array of elements markedly overlaps the *In re Wands* factors for undue experimentation employed when determining enablement<sup>1</sup>.

The guidelines instruct that each species claim should be analyzed to determine if either the entire structure is described or sufficient identifying characteristics are disclosed. For each genus claim an analogous determination is made regarding the presence of a representative number of species examples described either by complete structure or sufficient identifying characteristics. Again, validating a genus claim by evaluating a

---

<sup>1</sup> *In re Wands*, 8 USPQ2d 1400 (Fed. Cir. 1988)

representative number of species is analogous to the procedure used to determine enablement of a genus claim. See identification of this analogy recited in *University of California v. Eli Lilly and Co.*<sup>2</sup>

In this regard, the general overview provided by the interim guidelines represents well the relationship between written description and enablement. In particular, it reflects how the written description requirement is broader than the enablement requirement of Section 112, which is subsumed within the description of the invention and, thereby, satisfies a separate and distinct purpose in demonstrating possession of the invention.<sup>3</sup>

#### B) Genes, mRNA, and cDNA as Preamble Terms

The interim guidelines go to particular lengths to establish a distinction between two sets of preamble terms. One set consists of the terms "gene", "mRNA", and "cDNA". The PTO interprets each of these terms as representing a small genus of specific structures which include, in addition to the amino acid coding region, such elements as promoters, enhancers, and other regulatory elements. It is the PTO contention that all such subcomponents of these preamble terms must be described to satisfy the written description requirement.

This *ad hoc* interpretation establishes *per se* definitions of widely used molecular biology terms which are commonly used in patent law and molecular biology to mean different things depending on the particular context. Contrary to the interim guideline's interpretation, the most generally used context of these terms refers only to the coding portion of the molecules. This context is supported both in common patent usage and in case law.

Judge Rich provided an extensive background section on the molecular biology involved in protein synthesis in *In re O'Farrell*<sup>4</sup>. Nowhere does Judge Rich make the distinctions regarding the substructures suggested above. More recently, Judge Lourie provided an expanded background description of this topic in *In re Deuel*<sup>5</sup>. In that case, the claimed invention relates to isolated and purified DNA and cDNA molecules. While this case resolves an issue of obviousness, it derives its decision based upon analogous considerations regarding treating DNA/cDNA claims as chemical structures defined by their specific structure (e.g., sequence), rather than by their function or method of making. Needless to say, no distinction in cDNA substructure was given any consideration in evaluating the claims.

Finally, Example N: DNA in the previously mentioned "Training Materials for Examining Patent Applications with Respect to 35 U.S.C. Section 112, First Paragraph-

<sup>2</sup> *University of California v. Eli Lilly and Co.*, 43 USPQ2d 1398, 1406 (Fed. Cir. 1997)

<sup>3</sup> *Vas-Cath, Inc. v. Mahurkar*, 19 USPQ2d 1111 (Fed. Cir. 1991)

<sup>4</sup> *In re O'Farrell*, 7 USPQ2d 1673 (Fed. Cir. 1988)

<sup>5</sup> *In re Deuel*, 34 USPQ2d 1210 (Fed. Cir. 1995)

Enablement Chemical/Biotechnical Applications” describes and claims both DNA and cDNA molecules. The cDNA claims are in open “comprising” format reciting a specific nucleotide sequence or fragments of a specified length. It is clear from this common claim usage that the cDNA is intended to represent, and be synonymous with, the coding region of the molecule. Respondent is not aware of any patents where the substructure composition of cDNA, mRNA, or genes was a significant issue in determining aspects of patentability, or in determining what structures were deemed in the possession of the inventor.

Adoption of the PTO’s new definitions of cDNA, mRNA, and gene for purposes of written description considerations potentially could destabilize the economic infrastructure of the biotechnology community. Innumerable patents have issued claiming genes, cDNAs, and mRNAs without regard to the PTO’s new interpretation of claim language. Correspondingly, numerous business arrangements have been predicated upon such claims. Most, if not all, of those business deals would now be cloaked with uncertainty were these interim guidelines adopted. Such problems would far outweigh any benefit in waging a semantic debate over a *per se* definition of the structure of these molecules.

Representative of this curious parsing of molecular terminology, the interim guidelines provide a confusing interpretation of the following claim: “A gene comprising SEQ ID NO: 1”. This claim is described as being viewed as a species claim with a combination/subcombination relationship between the preamble and the body. In reality this is a genus claim based upon the open “comprising” transition phrase. This claim truly would be a species claim if redrafted using the transition phrase “consisting of”. Contrasting these two situations, it is clear that the genus/species nature of a claim is driven by the nature of the transition phrase and the body of the claim, not by the preamble.

If the body of the claim does not correspond well with the preamble term, this may represent a problem of definiteness under 35 USC 112, second paragraph; not the first paragraph of Section 112. A more appropriate consideration is if the “comprising” term enlarges the scope of the SEQ ID NO: 1 structure such that it is not enabled or does not support possession of the structure under the written description requirement.

### **C) Nucleic Acid, DNA, and RNA as Preamble Terms**

The interim guidelines propose that substitution of more general preamble terms, such as composition, nucleic acid, DNA, and RNA somehow creates a genus claim. The specific example of this phenomenon is the claim construction, “A nucleic acid comprising SEQ ID: 1.” The interim guidelines interpret the generic nature of this claim to reside in the term “nucleic acid.” Each member of the genus “nucleic acid” is considered under the interim guidelines to be a combination containing the subcombination “SEQ ID NO: 1” (which is a fragment of the nucleic acid). The interim guidelines proffer that the generic

nature of the term "nucleic acid" prevents a written description problem because one skilled in the art can readily envision a sufficient number of members of the claimed genus to provide written description support for the genus. A footnote "16" to pages 1405-1406 of the previously mentioned *University of California v. Eli Lilly and Co.* CAFC decision is recited to support this proposition.

Respondent has carefully reviewed this case, including the specified pages, and finds no mention of written description support for generic claims arising from envisioning genus members based on preamble terms such as composition, nucleic acid, DNA, or RNA. The issue addressed on pages 1405 to 1406 of this case is that a disclosure of rat insulin cDNA is not sufficient to support generic claims to vertebrates or mammals. Substitution of the term nucleic acid for cDNA does not remedy this deficiency. There is nothing in the term nucleic acid that envisions sufficient numbers of insulin sequences corresponding to different vertebrate or mammalian species so as to provide written description support for the genus. The only way to remedy the deficiency is to disclose the actual sequences of a representative number of species to support the genus; rather than wordsmith the preamble of the claim.

#### **D) What Defines Species Versus Genus Claims?**

The distinction between species and genus claims is an important concept in the interim guideline, because genus claims require additional considerations. At least in unpredictable arts such as chemical and biotechnology inventions, genus claims additionally require sufficient description of a representative number of species to support possession (written description) of the genus. The interim guidelines provide little direction and guidance toward distinguishing species from genus claims. As discussed above, where the interim guidelines address this issue between preamble terms such as cDNA and DNA, they appear to confuse the issue more than elucidate it. The generic nature or scope of a claim is determined by the interplay of the transition phrase and the limiting embodiments of the material representing the body of the claim. The following examples will attempt to illustrate this.

1) On page 1406 of *University of California v. Eli Lilly and Co.*, Judge Lourie sets forth a description of a classical chemical generic claim.

[i]n claims involving chemical materials, generic formulae usually indicate with specificity what the generic claims encompass. One skilled in the art can distinguish such a formula from others and can identify many of the species that the claims encompass. Accordingly, such a formula is normally an adequate description of the claimed genus.

In such a scenario, the generic formulae may be a three-ring heterocyclic nucleus of specified structure with two defined substituent R-groups (e.g., halogen and alkyl R-

groups) at specified locations on the nucleus. Species are envisioned or identified by substituting different members of each R-group (e.g., substituting a bromine or a chlorine as the halogen R-group). Enlarging the breadth of the claim by using "comprising" as the transition phrase permits inclusion of other unrelated compounds or materials (e.g., a solvent) without changing the generic formula. Such included unrelated compounds or materials need not be identified to satisfy the written description requirement. Similarly, choice of preamble phrases generally will not change the nature of the defined generic formula. For example, defining the preamble broadly as a composition, rather than a three-ring heterocyclic compound, does not permit adding a fourth ring or a third R-group to the defined nucleus. A person skilled in the heterocyclic art can distinguish this generic composition and species encompassed therein from others by its formula. Accordingly, the formula is an adequate written description of the claimed genus.

2) The first example in Section C(2) of the interim guidelines describes an isolated double-stranded DNA defined by sufficient identifying characteristics (i.e., size, cleavage map, and source from which the DNA is derived) that one skilled in the art would recognize from these characteristics that the inventor was in possession of the claimed material. This is a species claim regardless of the nature of the transition phrase associated with the claim. Changing the transition phrase from "consisting of" to "comprising" would broaden the scope of the claim by permitting additional unstated subject matter, but would not change the combination of characteristics that define this species of double stranded DNA.

3) The relationship described in the examples 1 and 2 above differs dramatically when the formula defining the invention in the body of the claim is a nucleic acid or amino acid sequence. In these cases, "consisting of" transitional language limits the claim to the recited nucleic acid or amino acid sequence. Substitution of "comprising" transitional language creates a generic sequence formula which permits additional unstated subject matter as previously. However, this open-ended language also permits the length of the nucleic acid or amino acid sequence to be expanded at either or at both ends. The magnitude of this lengthening of the original structure is indeterminate, and the identity of each added nucleotide or amino acid is unknown. The magnitude of this type of modification of the core sequence can be tempered by limiting the size of the claimed moiety (e.g., nucleic acid limited to 40 nucleotides), and by limiting the nature of the additional sequence (e.g., at a defined position in the amino acid sequence permit only lysines to be added to the carboxy end of the molecule).

The interim guidelines succinctly explain the inverse correlation between predictability in the art and the amount of disclosure necessary to satisfy the written description requirement. A generic formula must provide a reasonable expectation that species within that genus structure will exhibit similar function corresponding to the disclosed utility(ies) of the invention. Sufficient examples of species must be provided by the disclosure to support and validate possession (written description) and enablement of that level of predictability between structure and function. The broader the structure of the

genus and/or the more unpredictable the state of the art related to the invention, the more examples are required of the disclosure to establish that relationship. To better understand the dramatic consequences to the predictability of this structure-function relationship caused by comprising language in nucleic acid or amino acid sequence formulae, it may be instructive to analogize the effect of comparable modifications to more traditional generic chemical formulae.

In the context of example 1 above (generic three-ring heterocyclic formula), the analogous enhancement of claim scope would involve adding an indeterminate number of undefined rings to the nucleus of the molecule or adding an indeterminate number of new R-group substituents of undefined nature onto the nucleus. Such additional rings or R-groups would dramatically alter the structure-function relationship defining the claimed molecule. In other words, four-ring, seven-ring, or ninety-ring heterocyclic compounds would not be expected to exhibit the same function(s) or utility(ies) characteristic of a three-ring heterocyclic structure. Similarly, three-ring heterocyclic compounds having five, nine, or twelve substituent R-groups of undefined nature would not be expected to exhibit the same functions as species encompassed by the two defined R-groups of example 1. It is unlikely that any disclosure could support the possession or enablement of essentially an infinite array of possible structures in support of a real world patentable utility.

Examples A and B in the previously mentioned Training Materials for Enablement reinforce aspects of this marked enhancement of structural scope created through use of "comprising" transitional language in claims drawn to nucleic acid sequences. Both examples present related fact patterns involving claims reciting open "comprising" language and Markush groups containing specific Sequence ID numbers corresponding to three disclosed nucleotide sequences. Both examples rely upon hybridization involving the claimed sequences to effect the disclosed utility. Both examples cite a pair of literature references, Sambrook et al. and Wallace et al., for their teaching that mismatches within an oligonucleotide probe impart unpredictability to the hybridization process. Both examples explain how the "comprising" language markedly broadens the scope of the nucleic acid sequences by introducing additional random sequence of indeterminate length. In view of the teachings of Sambrook et al. and Wallace et al., the introduction of random base sequence was deemed to skew the predictability of structure to function sufficiently to render the claims nonenabled. Both examples recommended limiting the claim scope by using "consisting of" transitional claim language to satisfy the undue breadth problem.

The independent and distinct nature of the written description and the enablement requirements of Section 112, first paragraph contemplates situations where a chemical formula or a nucleic acid/amino acid sequence is described adequately so as to demonstrate possession at the time of filing, but that same disclosure fails to teach how to make or use (enable) the possessed invention. The distinctions between these two elements of Section 112 converge, however, where lack of enablement results from undue

breadth of claim structure (e.g., Examples A and B, above). Circumstances of undue breadth likely will invoke also a deficiency in the written description requirement. In both cases, there is failure to disclose sufficient numbers of species corresponding to the overly broad genus to support possession or enablement. The disclosure required for both possession and enablement is inversely correlated to similar levels of predictability in the art. Additional considerations drawn to the scope of claims, nature and field of invention, and level of skill in the art are similar when analyzing the same overly broad claims for possession and enablement.

This convergence of enablement and written description considerations is exemplified in *In re Fisher*<sup>6</sup>. Claim 4 of *Fisher* is drawn to an adrenocorticotrophic hormone (ACTH) preparation. The claim construction involves open-ended "containing" language (analogous to "comprising" language) with a limitation that the preparation is characterized as containing polypeptide of at least 24 amino acids having an enumerated sequence. The court indicated the open ended claim construction broadens the claim such that "the claimed subject matter is in no way limited by the presence, absence or sequence of amino acids beyond the 24<sup>th</sup> position." While this claim language was deemed definite under the second paragraph of Section 112, it raised questions of sufficiency of disclosure under the first paragraph of that section. The court ruled that the application failed to support ACTH preparations with other (i.e., greater) than 39 amino acids. Consequently, the specification was deemed to lack sufficient supporting description to comply with the requirements of 35 U.S.C. 112, first paragraph. Related to this lack of a sufficient supporting description, the claims were also found to be not enabled. This exemplifies the "broader" nature of the description requirement, wherein the lack of description was manifested also in the inability to make or obtain the invention.

The rationale for finding a written description deficiency in open-ended nucleotide/amino acid sequence claims, wherein the range of possible sequence structures far exceeds those taught or contemplated by the specification is in concert with a line of more recent decisions from the CAFC. Specifically, the decisions in *Amgen v. Chugai*<sup>7</sup> and *Fiers v. Sugano*<sup>8</sup>, as well as *University of California v. Eli Lilly and Co.* require a close correlation between defined sequence structure and the written description requirement.

Consistent with requiring precise and narrow disclosure of nucleic acid and amino acid sequences for purposes of written description and enablement, the CAFC also narrowly interprets nucleic acid structure considerations relative to determinations of obviousness. See *In re Bell*<sup>9</sup> and *In re Deuel*. Referring to *In re Bell* and *In re Deuel*, the CAFC in *University of California v. Eli Lilly and Co.* stated the following position on the relationship of written description of sequences to reaching conclusions of obviousness:

---

<sup>6</sup> *In re Fisher*, 166 USPQ 18,23 (CCPA 1970)

<sup>7</sup> *Amgen Inc. v. Chugai Pharmaceutical Co., Ltd.*, 18USPQ2d 1016 (Fed. Cir. 1991)

<sup>8</sup> *Fiers v. Sugano*, 25USPQ2d 1601 (Fed. Cir. 1993)

<sup>9</sup> *In re Bell*, 26 USPQ2d 1529 (Fed Cir. 1993)

[a] prior art disclosure of the amino acid sequence of a protein does not necessarily render particular DNA molecules encoding the protein obvious because the redundancy of the genetic code permits one to hypothesize an enormous number of DNA sequences coding for the protein. Thus, *a fortiori*, a description that does not render a claimed invention obvious does not sufficiently describe that invention for purposes of § 112, ¶ 1.

Thus, the court does not view a DNA sequence to be in the possession of an inventor for purposes of the written description requirement even when it is within the rubric of a known protein sequence and the array of codon correspondences defined by the genetic code. Consequently, it is most unlikely the court would consider the infinite population of possible nucleic acid sequences encompassed within the scope of open-ended "comprising" claim constructions to be in the possession of inventors based upon the disclosure of a fragmentary and minor subset of that population.

It should be noted that the above analysis regarding possession of a nucleic acid sequence based upon open-ended "comprising" claim language does not entail consideration of the breadth or nature of the preamble phrase of the claim. In view of the significance of the relationship between the transition phrase and the body of the claim (nucleotide or amino acid sequence) in determining the scope of the genus structure, considerations related to the preamble phrase are not controlling. The preoccupation of the interim guidelines with comparative analyses of different preamble phrases at the expense of the considerations enumerated above is misplaced, and makes the interim guidelines seriously deficient.

#### **E) Written Description Considerations Related to EST Claims**

EST product claims likely will be expressed in one of two major formats and in numerous formats of intermediate scope. Regardless of the format, the preamble will likely take multiple forms such as EST, cDNA, cDNA fragment, gene fragment, composition, DNA, DNA fragment, nucleic acid, polynucleotide, or probe. In its simplest form, the narrow scope format will be; "A [preamble phrase] consisting of SEQ ID NO: [ ]." This should represent a species claim falling within the "safe harbor" criteria described under Section C(1) of the interim guidelines, and the written description requirement would be satisfied.

The other major format represents a broad scope format. In its simplest form, the broad scope format will be, "A [preamble phrase] comprising SEQ ID NO: [ ]." This represents a genus claim of infinitely broad scope as there is no limitation on the number or sequence of nucleotides that may be added to the 5' or 3' ends of the disclosed SEQ ID NO: [ ] formula. Clearly, there will be a myriad of species that are not specifically described in the specification. The scope and level of unpredictability of the structure is so large that the person skilled in the art cannot envisage sufficient species to place the genus in possession of the inventor at the time of filing. The rationale for this conclusion

is based upon the examples and discussion developed above regarding application of the "comprising" transition phrase to claims drawn to nucleic acid or amino acid sequences.

Particular attention is directed to EST claims whose patentable utility is predicated upon the capacity of the EST sequence to function as a hybridization probe. Such utilities may include use of ESTs in forensic identification, tissue type or origin identification, chromosome mapping, chromosome identification, and tagging a gene of known and useful function. Dr. Alberts' communication of June 19, 1997 addressed considerations related to the enablement of such hybridization events. The examples and discussion presented in this response address additional considerations, such as mismatches within a nucleic acid probe, that impart unpredictability to the hybridization process. Specifically, reference is made to Examples A and B from the Training Materials for Enablement and the Sambrook et al. and Wallace et al. articles cited therein. The next article (48) in *Methods in Enzymology*, (Volume 152) after Wallace et al. is by William I. Wood, and is titled, "Gene Cloning Based on Long Oligonucleotide Probes"<sup>10</sup> (copy enclosed). On page 443, Wood states:

[o]nly probes of 17 or longer can be used to screen high-complexity libraries (e.g., a human genomic library). This is because the complexity of the mammalian genome is such that an exact match of any 16-base sequence would be expected at random. When a pool of sequences is used, the number of false positives can be a problem.

This teaching highlights serious problems related to broad EST genus claims reciting "comprising" as the transitional phrase. As indicated previously, such claims include the recited EST sequence corresponding to the SEQUENCE ID NO plus additional nucleic acid sequence attached to either or both ends of the molecule. This additional nucleic acid sequence is of indeterminate length and random sequence composition. Notwithstanding the specificity of the original SEQUENCE ID sequence corresponding to the disclosed EST, Wood teaches that additional overlapping sequences of at least 16 bases would hybridize randomly to regions throughout the genome. Random hybridization leads to false positives, and reduces the predictability of the EST claim structure relative to its disclosed function (utility). This random hybridization problem raises serious questions regarding the enablement of any disclosed utility that relies upon specific hybridization of the disclosed SEQUENCE ID structure.

From the perspective of the written description requirement, proportionately more examples of species sequences must be described in the specification as the size and unpredictability of the EST genus structure increases beyond the specific SEQUENCE ID structure. Since "comprising" transition language supports the introduction of an infinite amount of random sequence, such "comprising" genus claims will require a very large number of described species sequences to demonstrate possession of the claimed genus.

---

<sup>10</sup> W. I. Wood, *Methods in Enzymology* 152, 443-447 (1987).

The broad claim breadth discussed above is a function of the "comprising" transitional phrase and the body of the claim. The nature of the preamble phrase has an insignificant effect upon the nature of this claim scope. It is important to acknowledge this distinction regarding the contribution of the preamble phrase, because the interim guidelines are confusing in this regard. The interim guidelines appear to establish a *per se* rule that a genus claim in the format: "DNA, or nucleic acid, or composition comprising SEQUENCE ID NO: []" satisfies the written description requirement, because one skilled in the art can readily envision a sufficient number of members of the claimed genus. This representation regarding envisioning a sufficient number of members is recited to be related to the less specific, generic preamble language. There is no basis in fact or case law for this representation. By contrast, it is at odds with the controlling case law cited in this response. This representation is misleading, and draws attention away from the undue breadth of the genus claims, as well as the claim elements ("comprising" language in conjunction with the SEQUENCE ID NO: formula of the claim body) most responsible for establishing the breadth of the genus claims. In the Final Guidelines, the PTO is requested to address fully the relationship of written description to this claim scope issue. The treatment of this subject must address the relationship of "comprising" transition language to nucleic acid and amino acid sequences expressed as SEQUENCE ID NOs.

Thank you for the opportunity to present the views of the NIH. Please feel free to contact us, if we can be of further assistance.



Jack Spiegel, Ph.D  
Director, Division of  
Technology Transfer & Development  
Office of Technology Transfer  
National Institutes of Health  
(301) 496-7056 X289  
js45h@nih.gov

3 Enclosures



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service  
National Institutes of Health

Office of Technology  
National Institutes of Health  
6011 Executive Boulevard  
Rockville, MD 20852

March 21, 1997

Mr. Lawrence J. Goffney  
Deputy Assistant Secretary of Commerce and  
Deputy Commissioner of Patents and Trademarks

Dear Mr. Goffney:

This communication follows our telephone conversation of February 17, 1997 confirming your comments on expressed sequence tag (EST) patenting at the February 14th AAAS meeting in Seattle. We spoke at length of our mutual concerns that patent claims issuing on partial cDNA ESTs be commensurate in scope with their enabling disclosures. At the conclusion of our conversation, you kindly invited continued dialog with appropriate Group 1800 staff to advance this common goal. While there has been no communication back to me from Group 1800, debate on this issue continues to ferment within the biotechnology community at the National Institutes of Health (NIH), academic institutions, and industry. Therefore, I submit these written comments, herein, for consideration by you and your managers within the Biotechnology Group in anticipation that this formally initiates dialog between our offices on these matters of common interest.

Background - NIH Involvement in ESTs:

As you undoubtedly know, EST technology originated in the NIH laboratory of Dr. Craig Venter at the National Institute of Neurological Diseases and Stroke (NINDS), and we believe the first patent applications in this field were filed on Dr. Venter's discoveries by NIH in June 1991. A serious concern, at the time, was that public disclosure of EST sequences could create a prior art effect against subsequent patenting of newly discovered complete gene sequences possessing important diagnostic or therapeutic utilities. EST patent applications were filed, in significant measure, to provide short term insurance against such potential prior art blockage of future gene discoveries.

Contemporaneous with our invention filings, and continuing through the period of patent examination, the Court of Appeals for the Federal Circuit (CAFC) rendered a series of opinions drawn to DNA/protein sequences, including *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1016 (Fed.Cir.1991); *Fiers v. Sugano*, 25 USPQ2d 1601 (Fed.Cir.1993); *In re Bell*, 26 USPQ2d 1529 (Fed.Cir.1993); and subsequently, *In re Deuel*, 34 USPQ2d 1210 (Fed.Cir.1995) which reduced significantly our prior art concerns. These court decisions, in conjunction with the evolution of our definitive Patent Policy that moves away from patenting research tools, led to the NIH abandoning all pending EST applications. However, the NIH continues to actively pursue varied aspects of genomic research, including transferring the useful products of that endeavor to the private sector. Despite having no EST applications pending before your office, the NIH maintains an ongoing policy interest in these compositions as long as EST technologies potentially influence the development and availability of genomic inventions for the public health.

NIH Issues and Concerns:

Like many in the biotechnology research and intellectual property communities, the NIH is both surprised and disappointed by media reports indicating the USPTO now finds the use of ESTs as probes to be a specific utility in satisfaction of Section 101 requirements. Typically, the identity of the gene corresponding to an EST is not known. NIH scientists, as well as many in the academic community (i.e., those of ordinary skill in this art), view such bare EST sequence disclosures as providing little or no practical (real world) value toward advancing discovery in the human genome art. Potential value for EST sequences derives from future research relating EST sequences to genes or proteins of known function. At best, therefore, ESTs represent a poor research tool.

The relative merit of these discoveries as research tools should not militate against their public disclosure in any forum, including scientific literature, databases, or patents. The nature of such disclosures follows from the formats, policies, and statutes governing or established by each forum. We appreciate the unique character of the patent forum that balances the value of public disclosure against the right to exclude others from the claimed invention for a limited period of time. We appreciate also the mission of the USPTO to administer the patent statutes and rules to establish in each patent grant the appropriate *quid pro quo* between public disclosure and rights of exclusivity, thereby, advancing the constitutional mandate to promote the progress of science and the useful arts.

The interest of the NIH is to provide the PTO sufficient information to make appropriate and consistent decisions in establishing this *quid pro quo* relative to EST inventions. In this regard, NIH has concerns along two lines. First, we believe the utility of the typical EST invention may not meet the threshold criteria of utility set forth in 35 USC 101. Our second concern is more critical. If EST inventions do satisfy the specific utility requirements of patent law, we are concerned the PTO understands all the relevant issues to establish an appropriate balance between the value of EST disclosures and the scope of exclusionary claim protection. In other words, we are concerned how the PTO applies the enablement and description provisions of 35 USC 112 to establish a proper claim breadth for EST inventions.

As indicated previously, the NIH no longer has a proprietary position in patenting ESTs. Therefore, the NIH, similar to the PTO, has no interest in the commercial success or failure of any particular applicant or company in this arena. Rather, the NIH communicates our concerns in this regard, because we are sensitive to the possibility that prototypical ESTs with claim scope broad enough to encompass the corresponding complete gene sequence may unduly shift the intended *quid pro quo* in favor of the patentee. The systematic promulgation of this imbalance may have serious chilling consequences to further research and commercial development of diagnostic and therapeutic products related to human genomics. Clearly, this situation could negatively affect the public health, and the advancement of the public health is the mandate of this agency.

Our primary interest, in this regard, is that a new species of "submarine" patent not be spawned by unduly broad patents routinely issued in the EST art. This may arise through the congruence of two conditions. The first involves the expectation in the art that the PTO will issue EST patents routinely, wherein the utility of the corresponding gene is not known; yet claim scope is broad enough to encompass the entire gene. The second condition may develop as a consequence of the new sequence restriction practice, whereby up to ten nucleic or amino acid sequences may be

examined per application. The remaining sequences are withdrawn from consideration pending filing of a Divisional application. The net effect of this restriction process may be millions of confidential EST sequences lying in "limbo" within the PTO.

Upon subsequent and independent discovery of the complete gene, possessing a clinically and/or commercially significant utility, applicant resurrects the corresponding withdrawn EST species. The patent issuing on that Divisional potentially could block or retard development of the significant public health invention. The economics of this situation may be resolved within the marketplace for individual cases, and this should not be the preoccupation of the PTO or the NIH. However, a health care issue is created if industry delays or refrains from investing in this important endeavor because of rampant uncertainty surrounding the existence of submarine ESTs lurking within the PTO. This should be a concern not only to the NIH mandate, but also to the mandate of the PTO.

Patent procedures should not encourage submarine patents that undermine the pursuit of invention. In particular, submarine ESTs portray a singularly unsatisfactory perception of a secret disclosure of *de minimis* utility, whose only real function is to lay in predatory wait and feed off later developed inventions with significant health care utility. While it may fall within the four corners of legal patent prosecution procedure, this scenario does not advance the progress of science and the useful arts. The means by which the PTO can affect this process is to circumvent the perception that useless parasitic inventions are hibernating in the Central Files of the PTO. The public must feel confident the PTO will issue patents only for inventions with claims commensurate in scope with their specific "real world" utilities.

We believe the PTO can foster this public confidence by practicing consistently the guidelines regarding Section 101 and 112 issues it has already promulgated. What follows is intended to be the constructive input of a sister government agency with a common interest in advancing the progress of science and useful public health arts.

#### Legal Considerations regarding the Patenting of ESTs

I. The PTO should re-evaluate whether an asserted utility as a probe represents a "specific utility" for ESTs under 35 USC 101.

It is generally accepted that there is no practical utility in the use of a probe as an intermediate to analyze or make a final product gene of unknown function (utility). Indeed, most practitioners in this art believed the issue was resolved with release of the USPTO Utility Guidelines, along with its Legal Analysis Supporting Utility Examination Guidelines, and supporting documentation available through the USPTO Home Page.

One such supporting document, titled: "Synopsis of Application of Utility Guidelines With Examples" defines "Specific utility" in part as follows:

- a practical utility which defines a "real world" context of use.
- Utilities which require or constitute carrying out further research

to identify or reasonably confirm a "real world" context of use are not "specific utilities".

In the instant case, the context of use as a probe has no "real world" meaning until the gene for which the EST is a probe is identified, i.e., functionally characterized. A gene with no associated biological function has no "real world" meaning. Consequently, there can be no "specific utility" in probing a gene of unknown biological function. It requires further research to ascribe a biological function to an unknown gene probed by an EST in order to provide "real world" context. This deficiency in EST inventions follows from the fundamental Supreme Court ruling in *Brenner v. Manson*, 383 U.S. 519, 148 USPQ 689 that "a patent is not a hunting license. It is not a reward for the search, but compensation for its successful conclusion." This theme was advanced also by the Federal Circuit in *In re Brana*, 34 USPQ2d 1436 (Fed. Cir. 1995). Most recently, the Court of Appeals for the Federal Circuit in *Genentech v. Novo Nordisk*, Case 96-1440, decided March 13, 1997, invalidated a Genentech patent on cleavable fusion stating, "Genentech is attempting to bootstrap a vague statement of a problem into an enabling disclosure sufficient to dominate someone else's solution of the problem. This it cannot do."

The above indicated definition continues on to recite examples illustrating "specific utility", as well as examples illustrating the lack of "specific utility" resulting from a need for further research to identify or reasonably confirm a "real world" context of use. The instant EST scenario is consistent with three of the examples illustrating situations that do not define "specific utilities." One such example is "a method of assaying for or identifying a material that itself has no 'specific utility'". This is consistent with using an EST as a probe to assay or identify a gene that itself has no specific utility, because the gene has no known biological function. The second example is "a method of making a material that itself has no 'specific utility'". This is consistent with using an EST as a probe to make a gene that is itself has no specific utility, because the gene has no known biological function. The last example is most cogent and defines "a claim to an intermediate product for use in making a final product that has no known utility". Indeed, an EST probe for an unknown gene may be considered an intermediate for use in making a final product. In this case, the final product (i.e., the unknown gene) is a product with no known utility.

The last example regarding failure of the intermediate product to establish a "specific utility" is in concert with the probative case law. See *In re Joly*, 153 USPQ 45 (CCPA 1967) wherein Judge Rich instructs

\*\*\* the conclusion is inescapable that, just as the practical utility of the compound produced by a chemical process "is an essential element" in establishing patentability of the process, [Brenner v. Manson] 383 U.S. 519, 148 USPQ 689, so the practical utility of the compound, or compounds, produced from a chemical "intermediate," the "starting material" in such a process, is an essential element in establishing patentability of that intermediate. It seems clear that, if a process of producing a product of only conjectural use is not itself "useful", within Section 101, it cannot be said that the starting materials for such a process—i.e., the presently claimed intermediates—are "useful." It is not enough that the specification disclose that the intermediate exists and that it "works," reacts, or can be used to produce some intended product of no known use. Nor is it enough that the product disclosed to be

obtained from the intermediate belongs to some class of compounds which now is, or in the future might be, the subject of research to determine some specific use.\*\*\*

We conclude that appellants have not discharged their burden to show that the claimed subject matter is "useful" within the requirements of Section 101

Finally, following the "Synopsis of Application of Utility Guidelines" are a set of 12 examples of common biotechnology invention scenarios designed to walk the examiner/reader through the disclosure fact pattern, claims, and analysis of relevant utility issues. Example 9 is particularly applicable in being drawn to a disclosure and claims to a large number of cDNA fragments (ESTs). The disclosed utility of each EST is as a probe to obtain the corresponding full length gene. The full length gene is used prophetically to make the corresponding protein via routine recombinant methodologies. The protein product then is used to study cellular mechanisms and activities associated with the protein. Each EST is claimed individually using closed ("consisting of") language, wherein the claim is limited to the exact disclosed Sequence ID Number. The asserted utility for each EST is identified as a method of making the corresponding protein. The example goes on to explain that the probative determination of "specific utility" is whether or not the protein product has a "specific utility". Since the asserted utility for the protein was a research utility (not a "real world" or "specific utility under *Brenner v. Manson* criteria), the method of making that protein (the asserted utility of the claimed EST) necessarily could not define a "real world" context of use. The conclusion reached from the analysis was that no utility existed under 35 USC 101, and both Section 101 and Section 112, 1<sup>st</sup> paragraph enablement rejections are proper.

Therefore, the standard for "specific utility" of EST-probe intermediates is the objective factual determination whether the final gene or protein products themselves possess "real world" utility (i.e., known and defined biological function). Considerations regarding advances in the state of the art, the routine nature of hybridization and recombinant DNA techniques, or the relative efficiency of the probe-intermediate for its intended purpose are of little moment in this determination. Deviation from or failure to meet this standard leads invariably to a conclusion of no "specific utility" under 35 USC 101.

The NIH is not aware of any legislation or court decisions that negate or modify these reasoned guidelines recently developed by the PTO to analyze issues regarding "specific utility". It appears that a consistent application of those guidelines leads to the conclusion that EST probes have no "specific utility" under Section 101 if the corresponding gene is unknown or has no "specific utility."

## II If Deemed to Have Utility, The Scope of EST Claims Should Be Limited To The Specific EST Sequence.

It is well established that a deficiency under Section 101 creates also a deficiency under Section 112, first paragraph, since the specification cannot enable one skilled in the art to use an invention that is not useful. *Nelson v. Bowler*, 206 USPQ 881 (CCPA 1980). Therefore, ESTs failing the above "specific utility" test as probes to unknown genes should be considered not patentable also

under the enablement provision of 35 USC 112. In the event the PTO deems any or all ESTs satisfy the utility requirements for patentability, the following arguments are set forth regarding breadth of claim scope.

A. General Considerations Regarding Scope of Claims

While monitoring polynucleotide composition claims issued recently, we observe regularly that polynucleotide claims are broader than the disclosed Sequence ID Number listings. NIH observed open "comprising" language in conjunction with claim constructions, as well as the following forms of structural (sequence) variability: allelic variants; a polynucleotide complementary to a polynucleotide in a Markush group; and less than 100% sequence identity expressed as --at least "X" % (e.g., 95%) identity to a Markush group of polynucleotides.. It is not our intent to question the propriety of claim breadth in the particular fact situations of those issued patents. However, NIH has serious concerns if EST claims will issue with similar broad open-ended claim constructions. For the reasons developed in the sections to follow, we submit it should be a rare disclosure that supports EST claim scope broader than the specific Sequence ID Number.

B. PTO Guidelines/Training Materials for Examining Patent Applications With Respect to 35 USC Section 112, First Paragraph-Enablement Chemical/ Biotechnical Applications support the proposition that ESTs should be limited in scope.

On November 5, 1996, training material and guidelines analogous to the utility guidelines were made available to the public on the PTO Home Page. Included in the materials are two examples (Examples A & B) presenting related fact patterns and claims drawn to hybridization probes and methods of using same. Claims in the first example recite open "comprising" language and Markush groups containing specific Sequence ID numbers corresponding to three disclosed nucleotide probe sequences, ranging between 30 and 35 nucleotides in length, which hybridize specifically to a defined target of known utility. The second example differs by eliminating the functional limitation drawn to the specificity of hybridization. Both examples cite a pair of literature references, Sambrook et al. and Wallace et al., for their teaching that mismatches within an oligonucleotide probe impart unpredictability to the hybridization process.

In the analysis, both examples explain how the "comprising" language markedly broadens the scope of the probe embodiment by introducing random sequences of indeterminate length. In view of the teachings of Sambrook et al. and Wallace et al. regarding the effect of base mismatching on probe specificity, the introduction of random base sequences into a hybridization probe would require undue experimentation to identify or make all nucleic acid probes encompassed by and satisfying the functional (specificity) requirements of the claim. Therefore, hybridization-probe claims containing such open-ended "comprising" language would be subject to rejection under 35 USC 112, 1<sup>st</sup> paragraph. Both examples instruct replacing the open-ended "comprising" language with "consisting" language directed specifically to the disclosed probes in order to eliminate the undue breadth problem.

The second example indicates the need for functional claim limitations directed to the specificity and utility of the probe. The lack of such functional limitations exacerbates the Section 112, 1<sup>st</sup> paragraph deficiencies associated with "comprising" language by introducing additional enablement issues drawn to failure to teach how to use all the probes encompassed by the claim.

### C. Application of Enablement Guidelines to EST Inventions

The hybridization-probe examples in the PTO Enablement Guidelines relate to EST cases asserting a probe utility. EST sequences, corresponding to unknown genes or genes of unknown function (specific utility), relate to at least Example 2, above. Regardless the breadth of claim language, we submit this scenario fails the "how to use" considerations of enablement under Section 112, 1<sup>st</sup> paragraph by teaching the use of an EST moiety to probe specifically for an unknown structure. If a probe does not bind specifically, how can it be distinguished from other probes? If the EST does bind something, the person skilled in the art would not know if the probe bound the intended species. It is left to experimentation outside the teachings of the specification to define all the parameters of a successful hybridization and, thereby, the real use of the claimed probe. It is not routine in this art to require the user to discover new genes in order to use a patented hybridization probe. At the very least, this would fall into the category of requiring undue experimentation.

EST sequences related to genes of known function are not subject to the "specific utility" criticisms described previously, but do correspond to the scenario outlined in the first hybridization example, above. Consequently, EST product and method of use claims with closed "consisting of" language drawn to specific Sequence ID Numbers should be free of Section 112 enablement criticisms, assuming an adequate written disclosure setting forth how to make and use the invention, including the best mode. Introducing open "comprising" claim language, of course, would trigger the above indicated "Undue Breadth" rejection under Section 112, 1<sup>st</sup> paragraph. The introduction of more moderate claim broadening language, however, such as allelic variants, fragments thereof, having at least 90% identity; a polynucleotide complementary to, etc., should be analyzed for undue breadth on a claim by claim basis using standard *Ex parte Forman / In re Wands* considerations viewed from the perspectives elucidated in Example 1, above.

An additional factor to be considered when ESTs are used to probe genomic DNA is that peptide coding regions (exons) generally are interrupted by non-coding introns. EST probes, derived from cDNA, reflect only exon and regulatory sequences from the genomic polynucleotide population. Consequently, the nucleotide sequence of EST probes may not be contiguous with the corresponding genomic DNA. Indeed, an EST sequence may have homology to several discontinuous regions of genomic DNA separated by multiple exons. Under such circumstances, only a variable fraction of each EST sequence probe would actually hybridize to the gene. Depending on a number of factors, including the size of the EST, introducing variability into such EST structure via claim broadening language may compromise the ability of the EST to function as a probe in concert with the considerations set forth in the Sambrook et al. and Wallace et al. citations discussed in Example 1, above. It would be expected, therefore, that any claim broadening language would be supported by appropriate working examples addressing these issues. In this regard, information gained from one EST species does carry over to different EST species. This should necessitate different working examples for each claimed EST species.

Mr. Lawrence J. Goffney  
March 21, 1997  
Page 8

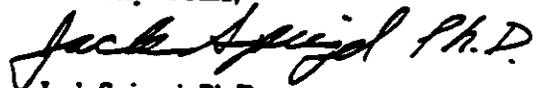
D. Additional Considerations

As indicated previously, "comprising" claim language encompasses additional random DNA sequences different from those specifically disclosed in the application. Applicants cannot describe or envisage the structure of these additional sequences. Consequently, the specification must be defective under the "description" requirement of Section 112, 1<sup>st</sup> paragraph. This interpretation is in concert with a line of Federal Circuit decisions involving nucleic acid and amino acid structures. See *Fiers v. Sugano*, 25 USPQ2d 1601 (Fed.Cir.1993) and *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1016 (Fed.Cir.1991). Also see *In re Bell*, 26 USPQ2d 1529 (Fed.Cir.1993) and *In re Deuel*, 34 USPQ2d 1210 (Fed.Cir.1995) regarding related issues drawn to treating nucleic acids as chemical structures. Furthermore, the Examiner cannot search the prior art to determine novelty and unobviousness without knowledge of which sequences are being claimed. A claim that does not define the invention with sufficient distinctness to permit a proper search of the prior art is invalid also under the second paragraph of 35 USC 112.

In summary, the NIH believes EST sequences for use as probes do not satisfy the utility requirements under Section 101 unless the EST sequences correspond to genes of known function. Furthermore, the scope of all EST probe claims should be limited to "consisting of" language consistent with the PTO guidelines for enablement issues. Indeed, we caution against any broadening of the scope of EST claims beyond the disclosed Sequence ID Number. We feel a consequence of issuing broad claims to EST sequences as probes to unknown genes could be the emergence of "submarine" patents having a chilling effect on development of genomic products for the public health. The solution to this potential health care problem merely requires the PTO to strictly and consistently adhere to existing utility and enablement examination guidelines established during the last two years.

I appreciate this opportunity to present the views of the NIH. I again hope this communication initiates ongoing dialog between our offices to advance these issues. Please feel free to contact me if I can be of any assistance.

Sincerely Yours,



Jack Spiegel, Ph.D.  
Director, Division of Technology  
Development and Transfer  
Office of Technology Transfer  
National Institutes of Health  
Phone: (301) 496-7056 X289  
FAX: (301) 402-0220  
E-Mail: js45h@nih.gov

cc: Dr. Harold E. Varmus  
Dr. Michael M. Gottesman  
Dr. Maria C. Freire

*Methods in Enzymology*

*Volume 152*

*Guide to Molecular Cloning  
Techniques*

EDITED BY

*Shelby L. Berger*

NATIONAL CANCER INSTITUTE  
NATIONAL INSTITUTES OF HEALTH  
BETHESDA, MARYLAND

*Alan R. Kimmel*

NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES  
NATIONAL INSTITUTES OF HEALTH  
BETHESDA, MARYLAND

1987



ACADEMIC PRESS, INC.

Harcourt Brace Jovanovich, Publishers

Orlando San Diego New York Austin  
Boston London Sydney Tokyo Toronto

wrapped with plastic between two Cronex me (7)).

s)

electrophoresis of d for 30 min in 0.1 bided for the DNA (NA) appear to be are a described for

s)

ents less than 1.5 ization membrane om the agarose gel s. Acid depurina- fer out of the gel. idize efficiently to esis of DNA and ne [8. 61]. ne conditions de-

collaboration of Dr. RBW is a member of

S.A. 76, 3683 (1979).

## [48] Gene Cloning Based on Long Oligonucleotide Probes

By WILLIAM I. WOOD

The most commonly used technique for gene cloning has been to utilize oligonucleotide probes based on protein sequence data. Of course this approach requires characterized and purified protein so that at least a portion of amino acid sequence can be determined and used to infer the corresponding DNA sequence. Based on the amino acid sequence information, either short or long oligonucleotide probes can be synthesized chemically.

Short probes are typically 11–20 bases in length and are pools of 8–32 (or more) sequences including all of the possible codon choices for each amino acid. There are three disadvantages of short probes. (1) They can generally only be used in regions of low codon redundancy; otherwise the pool size becomes unmanageable. (2) The amino acid sequence must be correct. A single mismatch is generally sufficient to prevent hybridization of the probe. (3) Only probes of 17 or longer can be used to screen high-complexity libraries (e.g., a human genomic library). This is because the complexity of the mammalian genome is such that an exact match of any 16-base sequence would be expected at random. When a pool of sequences is used, the number of false positives can be a problem. In some cases this difficulty can be overcome by using two nearby short probes. The advantage of short probes is that if the protein sequence data are correct, the probe should hybridize faithfully as all the codon choices are covered. Also, the exact hybridization conditions used need not be determined empirically when tetramethylammonium chloride is used<sup>1</sup> (see also this volume [49]).

Long probes on the other hand are typically 30–100 nucleotides long and are a single sequence based on a best guess for each codon. The long probe approach was first used to screen for three different genes: bovine trypsin inhibitor,<sup>2</sup> human insulin-like growth factor I,<sup>3</sup> and human factor IX.<sup>4</sup> There are three advantages of long probes. (1) Any stretch of amino acid sequence 10 or longer can be used; regions of low redundancy while

<sup>1</sup> W. I. Wood, J. Gitschier, L. A. Lasky, and R. M. Lawn, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1585 (1985).

<sup>2</sup> S. Anderson and I. B. Kingston, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6838 (1983).

<sup>3</sup> A. Ulrich, C. H. Berman, T. J. Dull, A. Gray, and J. M. Lee, *EMBO J.* **3**, 361 (1984).

<sup>4</sup> M. Jaye, H. de la Salle, F. Schamber, A. Ballard, V. Kohij, A. Findeli, P. Tolstoshev, and J. P. Lecocq, *Nucleic Acids Res.* **11**, 2325 (1983).

always a help are not especially important. (2) The amino acid sequence need not be absolutely correct. An erroneous amino acid or two can be tolerated.<sup>5,6</sup> (3) These probes can be used to screen high-complexity libraries with fewer false positives. The only disadvantage of long probes is the uncertainty of the codon choice. With the right codon choice even a 30-mer will hybridize very specifically in a high complexity library screen. However, if the codon choices are completely incorrect, the probe will never hybridize under any conditions.

In spite of the uncertainties over codon selection, the long probe approach is currently the method of choice in screening for genes based on protein sequence data. This is not to say that pools of short oligonucleotides do not have utility. The wisest course in screening for any new gene is to pursue all the avenues possible consistent with the available time, energy, and manpower. However, the utility of the long, single sequence probes has been demonstrated repeatedly for the screening of high-complexity libraries starting with any stretch of protein sequence data.

A variety of codon usage information can be used depending on the particular gene to be screened. A number of workers have used the codon usage table for mammalian genes compiled by Grantham *et al.*<sup>7</sup> It is also possible to use the codon frequency of a gene already cloned from the same or a related family. The most extensive consideration of the codon usage problem is given by Lathe<sup>8</sup> which includes considerations to reduce the number of CG base pairs in adjacent codons and other optimizations. It should also be noted that some workers have used long probes which are pools of 8 or 16 sequences covering several common codon choices at a few of the amino acids.<sup>9,10</sup> In some respects this approach encompasses some of the best features of the long and short probe approaches.

The probe can be labeled by any of the standard techniques including end labeling, filling in with the Klenow fragment of DNA polymerase I (used with two long oligonucleotides that have 12- to 15-base overlap), cloning in M13 and primer extending across the probe region, and even

<sup>5</sup> A. Ullrich, J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petruzelli, T. J. Dull, A. Gray, L. Coussens, Y.-C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran, *Nature (London)* 313, 756 (1985).

<sup>6</sup> D. Pennica, G. E. Nedwin, J. S. Hayflick, P. H. Seeburg, R. Derynck, M. A. Palladino, W. J. Kohr, B. B. Aggarwal, and D. Goeddel, *Nature (London)* 312, 724 (1984).

<sup>7</sup> R. Grantham, C. Garter, M. Gouy, M. Jacobzone, and R. Mercier, *Nucleic Acids Res.* 9, 143 (1981).

<sup>8</sup> R. Lathe, *J. Mol. Biol.* 183, 1 (1985).

<sup>9</sup> P. H. Seeburg and J. P. Adelman, *Nature (London)* 311, 666 (1984).

<sup>10</sup> J. J. Toole, J. L. Knopf, J. M. Wozney, L. A. Sultzman, J. L. Buecker, D. D. Pittman, R. J. Kaufman, E. Brown, C. Shoemaker, E. C. Orr, G. W. Amphlett, W. Barry Foster, M. L. Coe, G. J. Knutson, D. N. Fass, and R. M. Hewick, *Nature (London)* 312, 342 (1984).

no acid sequence  
 id or two can be  
 gh-complexity li-  
 : of long probes is  
 on choice even a  
 ity library screen.  
 ct. the probe will

ne long probe ap-  
 or genes based on  
 of short oligonu-  
 ning for any new  
 with the available  
 the long, single  
 the screening of  
 protein sequence

depending on the  
 ve used the codon  
 7 *et al.* It is also  
 d from the  
 tion of the codon  
 erations to reduce  
 ner optimizations.  
 ong probes which  
 codon choices at  
 ach encompasses  
 approaches.  
 hniques including  
 NA polymerase I  
 15-base overlap),  
 region, and even

E. J. Dull, A. Gray, L.  
 g. C. Grunfeld, O. M.

rnck, M. A. Palladino,  
 12, 724 (1984).

. *Nucleic Acids Res.* 9,

4).

jecker, D. D. Pittman,  
 ilet, W. Barry Foster,  
*ure (London)* 312, 342

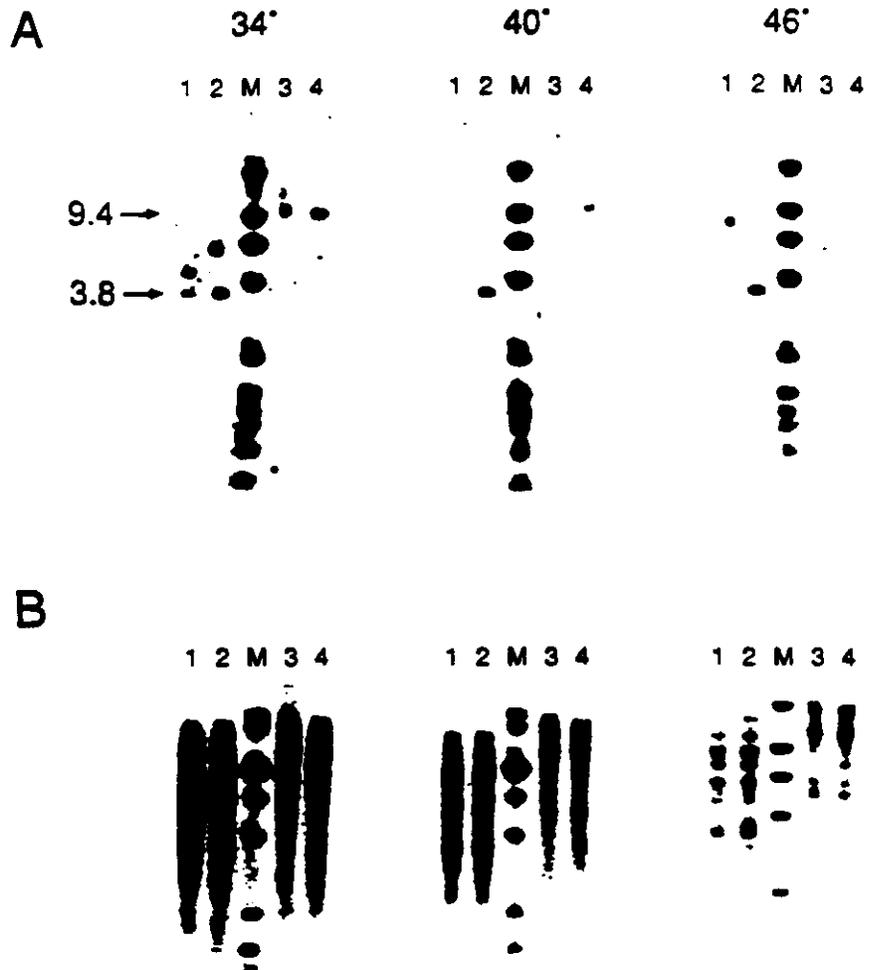


FIG. 1. Long probe hybridization to genomic blots. Two long probes for human factor VIII were hybridized to genomic blots of DNA, containing 1 and 4 X chromosomes.<sup>11</sup> Since factor VIII was known to be on the X chromosome, authentic hybridization is indicated by a relative intensity of 1:4 rather than 1:1 for the autosomes. Human genomic DNA, 5  $\mu$ g, was digested with *EcoRI* or *BamHI*, separated electrophoretically in an agarose gel, and transferred to nitrocellulose. Lane 1, *EcoRI* male DNA; lane 2, *EcoRI* 46.XXXXY DNA; lane M, markers, end-labeled  $\lambda$ HindIII and  $\phi$ X174 HaeIII; lane 3, *BamHI* male DNA; lane 4, *BamHI* 46.XXXXY DNA. The hybridization was as described in the text. The blots were washed in  $1 \times$  SSC, 0.1% SDS at the temperature indicated. (A) A 36-mer end-labeled probe; (B) an 81-mer probe, cloned in M13 and labeled by fill-in reaction.

TABLE I  
SELECTED LIST OF GENES ISOLATED WITH LONG PROBES

Protein	Oligonucleotide length (bases)	Match*	Reference
Trypsin inhibitor	86	...5x11x8...	2
Insulin-like growth factor I	94	...5x11x2x5x8x5...	3
Factor IX	52	...14x5x12	4
Transforming growth factor $\alpha$	74	...14x2x6...	12
Luteinizing-hormone-releasing hormone	38	...14...	9
Tumor necrosis factor $\alpha$	42	...8x17...	6
Factor VIII	36	...14x2x10...	11
Insulin receptor $\alpha$	63	11x8x11x3x10...	5
Insulin receptor $\beta$	54	...5x12...	5
$\beta$ -Adrenergic receptor	38	14x3x19	13

\* Numbers give the bases in a row which match. Each mismatch is one x. The three dots (...) indicate an extension of the probe with poor match.

end labeling a series of short overlapping oligonucleotides and ligating them together to make a long probe.<sup>3</sup> (See this volume [10] for labeling methods.)

The labeled probes can be used to screen any available cDNA or genomic library in a plasmid or phage vector. The most common current practice would be to screen a cDNA library in  $\lambda$ gt10. Depending on the strength of the hybridization, it is best to screen about 10,000 to 20,000 phage on each 150-mm plate.

Commonly used screening conditions are to hybridize in  $6\times$  SSC, 50 mM sodium phosphate (pH 6.8),  $5\times$  Denhardt's solution, 0.1 g/liter boiled, sonicated salmon sperm DNA, 20% formamide, and 10% dextran sulfate at 42° and to wash in  $0.2\times$  SSC, 0.1% SDS at 37°.<sup>3</sup> Prior to screening a library for the gene, it is often informative to hybridize the probes to genomic blots to determine whether a unique band can be observed. These blots can be used for two purposes: (1) to test different wash conditions to find the most stringent wash possible (see also this volume [43,45,61]) and (2) to determine which of several long probes is the most likely to be useful, thus eliminating some probes as having insufficient specificity to be of further use (especially in screening genomic libraries). Figure 1 shows the hybridization of two oligonucleotide probes to human *EcoRI*- and *BamHI*-digested DNA.<sup>11</sup> The probes were hybridized under

<sup>11</sup> W. I. Wood, D. J. Capon, C. C. Simonsen, D. L. Eaton, J. Gitschier, B. Keyt, P. H. Seeburg, D. H. Smith, P. Hollingshead, K. L. Wion, E. Delwart, E. G. D. Tuddenham, G. A. Vehar, and R. M. Lawn, *Nature (London)* 312, 330 (1984).

PROBES	Reference
	2
5x5...	3
	4
	12
	9
	6
	11
...	5
	5
	13

match is one x. The  
or match.

... and ligating  
... for labeling

available cDNA or  
most common current  
. Depending on the  
at 10,000 to 20,000

dize in 6x SSC, 50  
olution. 0.1 g/liter  
e, and 10% dextran  
7°. <sup>3</sup> Prior to screen-  
ridize the probes to  
l can be observed.  
test different wash  
ee also this volume  
; probes is the most  
having insufficient  
; genomic libraries).  
de probes to human  
re hybridized under

tschier, B. Keyt, P. H.  
t. E. G. D. Tuddenham,  
J.

the conditions given above and washed in 1x SSC, 0.1% SDS at 34°, 40°, and 46°. Probe A, a 36-mer, clearly is much more suitable than probe B, an 81-mer, from another portion of the same protein. A screen of a genomic library with probe A isolated the derived clones easily. A screen with probe B was not attempted because of the large number of hybridizing bands. Another example of genomic blots with long probes has been published.<sup>2</sup>

Table 1 lists several examples where long probes based on protein sequence data have been used to clone new genes.<sup>12,13</sup> The table is not intended to be complete but purports to show the kinds of sequence match necessary for suitable hybridization. Of primary importance in obtaining adequate hybridization is the number of nucleotides in a row which match rather than the percentage homology, although clearly two regions separated by a single mismatch hybridize better than the longer of the two regions alone.

<sup>12</sup> R. Derynck, A. B. Roberts, M. E. Winkler, E. Y. Chen, and D. V. Goeddel. *Cell (Cambridge, Mass.)* 38, 287 (1984).

<sup>13</sup> R. A. F. Dixon, B. K. Koblika, D. J. Strader, J. L. Benovic, H. G. Dohman, J. Frielle, M. A. Bolanowski, C. D. Bennett, E. Rands, R. E. Diehl, R. A. Mumford, E. E. Slater, I. S. Sigal, M. G. Caron, R. J. Lefkowitz, and C. D. Strader. *Nature (London)* 321, 75 (1986).

#### [49] Hybridization of Genomic DNA to Oligonucleotide Probes in the Presence of Tetramethylammonium Chloride

By ANTHONY G. DiLELLA and SAVIO L. C. WOO

In this chapter we present a powerful method for the hybridization of genomic DNA to AT-rich oligonucleotide probes. The method utilizes tetramethylammonium chloride (TMA), a reagent which binds AT-rich DNA polymers<sup>1</sup> while concomitantly abolishing the preferential melting of AT versus GC base pairs.<sup>2</sup> Thus, hybridization of these probes to genomic DNA (Southern blots) or to recombinant DNA libraries is a function of probe length; it occurs in a manner independent of base composition.<sup>3</sup> The technique is particularly well suited for detecting and dis-

<sup>1</sup> J. T. Shapiro, B. S. Stannard, and G. Feisenfeld. *Biochemistry* 8, 3233 (1969).

<sup>2</sup> W. B. Melchior and P. H. von Hippel. *Proc. Natl. Acad. Sci. U.S.A.* 70, 298 (1973).

<sup>3</sup> W. I. Wood, J. Gitschier, L. A. Lasky, and R. M. Lawn. *Proc. Natl. Acad. Sci. U.S.A.* 82, 1585 (1985).