REFLECTIONS ON THE SCIENCE AND LAW OF STRUCTURAL BIOLOGY, GENOMICS, AND DRUG DEVELOPMENT

Helen M. Berman*
Rochelle C. Dreyfuss**

The Patent Act is now over a half-century old, and many observers have become concerned that it is not responsive to the needs of emerging industries or to the changing landscape of science. In this Article, we look at this issue in the context of the life sciences and examine how patent doctrine has reacted as the fields of proteomics, genomics, and structural biology have advanced. We find many missteps along what we call the subject matter-inventiveness-scope-exemption trajectory. Patents now protect subject matter such as genes and proteins. These advances no longer represent particularly difficult scientific challenges. But because this subject matter lies far upstream, the patents on these advances have strong potential to block drug development. Although attempts have been made to narrow the scope of upstream patents to permit competitive pharmaceutical research, patentees have fought back with new claiming strategies and with new arguments for curtailing use of their inventions for experimental purposes.

In this Article, we argue that a more realistic appraisal of the underlying science is needed, one that takes account of the automation of early stage research and recognizes that the place for patenting is downstream—when long, convoluted, and risky creative efforts are made to convert genetic and proteomic knowledge into viable commercial products. Pushing patents downstream is not only more defensible from the perspective of the underlying science, it is also desirable socially, for it would increase public access to the fundamental building blocks of scientific knowledge and allow competitive basic research to flourish.

---

* Board of Governors Professor of Chemistry and Chemical Biology, Rutgers University; Director, Protein Data Bank, Research Collaboratory for Structural Bioinformatics.

** Pauline Newman Professor of Law, New York University School of Law. We wish to thank Melissa Feeney Wasserman, NYU Class of 2007, for her extraordinary work on this Article. We are also grateful to Eileen Kane and Marty Adelman for their comments, Nicole Arzt for technical support, and the Filomen D'Agostino and Max E. Greenberg Research Fund for financial support.

Genomics,[1] proteomics,[2] and structural biology[3] are relatively young sciences. Their roots lie in the discovery of the double-helix structure of DNA in 1953,[4] in the first use of X-ray crystallography to determine the three-dimensional structure of proteins in 1957,[5] and in the development of recombinant DNA technology in the 1970s.[6] The advances in these disciplines have allowed scientists to explore biological properties at the molecular level, to isolate individual genes, and to discover the composition and structure of the proteins that each gene instructs the organism to express. This information establishes a structural and rational basis with which to search for substances (drugs) that affect the functioning of molecules in their cellular environments. In its 1980 decision in *Diamond v. Chakrabarty*,[7] the U.S. Supreme Court assured that work in genomics, proteomics, and

1. Genomics is the comprehensive study of the genetic information of a cell or organism.

2. Proteomics is the large-scale study of proteins, with emphasis on their structures and functions.

3. Structural biology is the study of the three-dimensional structures of biological macromolecules, such as proteins and DNA, with the goal of understanding their functions.

4. J.D. Watson & F.H.C. Crick, *Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid*, 171 NATURE 737 (1953).

5. J.C. Kendrew et al., *A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-ray Analysis*, 181 NATURE 662 (1958).

6. Paul Berg et al., *Letter: Potential Biohazards of Recombinant DNA Molecules*, 185 SCIENCE 303 (1974); Stanley N. Cohen et al., *Construction of Biologically Functional Bacterial Plasmids In Vitro*, 70 PROC. NAT'L ACAD. SCI. 3240 (1973); David A. Jackson et al., *Biochemical Method for Inserting New Genetic Information Into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of* Escherichia Coli, 69 PROC. NAT'L ACAD. SCI. 2904 (1972); U.S. Patent No. 4,237,224 (filed Jan. 4, 1979).

7. 447 U.S. 303 (1980). For the commercial interest in the decision, see Rebecca S. Eisenberg, *The Story of* Diamond v. Chakrabarty*: Technological Change and the Subject Matter Boundaries of the Patent System*, *in* INTELLECTUAL PROPERTY STORIES 327, 342–43 (Jane C. Ginsburg & Rochelle Cooper Dreyfuss eds., 2006).

structural biology would attract commercial financing by holding that a genetically engineered microorganism is patentable subject matter.

In the quarter century since that decision, the biotechnological sciences have advanced significantly; however, the law on patenting biotechnological inventions has evolved much more slowly. *Chakrabarty* essentially changed the default position on protecting life-sciences materials. Its broad holding—that the subject matter of patent law extends to "anything under the sun made by man"[8]—means that developments in these fields are now presumed to be patentable. As a result, Congress has had little occasion to intervene, or even consider, patent law issues regarding genomics or proteomics.[9] And even Supreme Court involvement has become minimal, for at around the same time that biotechnological research began to flourish, Congress established the Court of Appeals for the Federal Circuit.[10] The near-monopoly that this court exerts over patent appeals ended the Supreme Court's need to hear patent cases in order to resolve circuit conflicts in patent jurisprudence.[11]

In this Article, we argue that it is time for this "post-*Chakrabarty* default" to end, and that a systematic reevaluation of how patent law applies to genes, proteins, and related inventions must begin. Not only have legal developments pertaining to these advances in science and technology escaped scrutiny at the highest levels, these developments present an issue that is somewhat new to patent law: Because much of the protected material is unique, the patents that are issuing cannot always be avoided by inventing a substitute.[12] We take as our starting point the position Dan Burk and Mark Lemley have so persuasively articulated, that patent doctrine is not as

8. *Chakrabarty*, 447 U.S. at 309.

9. *See* Eisenberg, *supra* note 7, at 357. Congress added a new paragraph to 35 U.S.C. § 103, but has otherwise done very little to tailor patent law specifically to biotechnology. *See* 35 U.S.C. § 103(c) (2000 & Supp. 2005).

10. Federal Courts Improvement Act of 1982, Pub. L. No. 97-164, 96 Stat. 25 (codified as amended in scattered sections of 28 U.S.C.).

11. *See generally* Rochelle Cooper Dreyfuss, *The Federal Circuit: A Case Study in Specialized Courts*, 64 N.Y.U. L. REV. 1 (1989); John F. Duffy, *The* Festo *Decision and the Return of the Supreme Court to the Bar of Patents*, 2002 SUP. CT. REV. 273; Mark D. Janis, *Patent Law in the Age of the Invisible Supreme Court*, 2001 U. ILL. L. REV. 387.

12. In a recent piece, Mark Lemley observes that nanotechnology presents several new challenges to patent law, including the recognition of rights on ideas and on inventions that cross industrial sectors, as well as increased participation in patenting by universities. While this is surely correct, it can also be said that these problems were foreshadowed by biotechnology. *See* Mark A. Lemley, *Patenting Nanotechnology* (Stanford Law & Econ. Olin Working Paper No. 304, 2005), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=741326; *see also* Rochelle Cooper Dreyfuss, *Unique Works/Unique Challenges at the Intellectual Property/Competition Law Interface*, *in* EUROPEAN COMPETITION LAW ANNUAL 2005—THE INTERACTION BETWEEN COMPETITION LAW AND INTELLECTUAL PROPERTY LAW (Claus-Dieter Ehlermann & Isabela Atanasiu eds., forthcoming), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=763688.

transsubstantive as might appear, but rather, that it is "industry specific" and its gloss is shaped by the disciplines to which it is applied.[13] But even if one regards patent law as transsubstantive, it can be argued that it is time to bring the law into alignment with the altering face of science. Genomics, proteomics, and structural biology furnish a context in which to investigate the sorts of changes that are needed.[14]

Thus, we begin in Part I with a description of how biotech patent law has been informed by a particular perception of rational drug development. In Part II, we move on to examine how work in that field is actually conducted and organized. In Part III, we show how these accounts differ and then ask whether—and how—patent law should change to promote genomic and proteomic research more effectively.[15]

I. THE MYTH OF RATIONAL DRUG DEVELOPMENT AND ITS EFFECT ON PATENT LAW

The narrative on which biotechnology patent law is based is easily described: It is founded on the principle that there are straightforward relationships between genes and the proteins that they encode, and between the geometric structures of proteins and potential pharmacological agents. As one commentator, summarizing the hopes of early scientists, put it:

> In many situations, genomics should reduce the preclinical research time necessary to identify the relevant molecular "target" for a given disease. Consider the situation of a researcher who is interested in a particular disease. She may be able to use expression profiling/DNA chip technology to identify the gene sequence that is overexpressed in the cells of persons with that disease. Having identified the relevant DNA sequence, she should then be able to use other bioinformatic techniques to determine, in relatively short

13. *See* Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?*, 17 BERKELEY TECH. L.J. 1155 (2002); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575 (2003).

14. This Article evolved from discussions between the authors while we served on a Committee on Intellectual Property in Genomic and Protein Research and Innovation, which was convened under the auspices of the National Academies of Science. The recommendations in this Article are, however, strictly our own. For the Committee report, see NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., REAPING THE BENEFITS OF GENOMIC AND PROTEOMIC RESEARCH: INTELLECTUAL PROPERTY RIGHTS, INNOVATION, AND PUBLIC HEALTH (2006) [hereinafter NAS 2006 REPORT], which is available at http://books.nap.edu/catalog/11487.html.

15. For an account of the advances in gene sequencing, see Judge Patti Saris's explanation of the research involved in *Amgen, Inc. v. Chugai Pharmaceutical Co.*, No. 87-2617-4, 1989 WL 169006 (D. Mass. Dec. 11, 1989), *aff'd in part, vacated in part*, 927 F.2d 1200 (Fed. Cir. 1991) (regarding cloning the gene for erythropoietin (EPO)).

> order, the function of the protein encoded by the gene and whether this protein could be an appropriate target for drug development. She will be able to do so by inputting the human DNA into a database that looks for similarities, or homologies, between the human DNA and previously sequenced DNA segments (either human DNA or the DNA of other organisms). If a match or near match is found, the researcher will be able to make an educated guess as to the function of the protein and as to whether it will be an appropriate target [for testing the therapeutic effects of particular drug candidates].[16]

This optimistic view of drug discovery in the genomic era was of critical importance to early thinking about the application of patent law to this field, because it made it appear that the crucial advance in the development of every new drug was (or was soon to be) the association of a new gene or protein sequence with a particular disease. Since everything else would, legally speaking, flow obviously from that crucial step, the need for patents to support drug development quickly translated into a legal regime that awards patent rights to genetic and proteomic materials at an early stage, when sequencing (and correlation with disease) first occur.

Granting patents on gene sequences quickly had a cascading effect along what might be called the "subject matter-inventiveness-scope-exemption trajectory." Because there is a finite—and surprisingly small—number of genes, it was soon recognized that each gene codes for many proteins and takes on different regulatory roles depending on the environment in which it operates.[17] For the law, this means—as Eileen Kane so neatly put it—that one gene patent could potentially generate a "molecular portfolio" of rights—rights over all of the derivative molecules generated by the gene, as well as all of their functions.[18]

---

16. Arti K. Rai, *The Information Revolution Reaches Pharmaceuticals: Balancing Innovation Incentives, Cost, and Access in the Post-Genomics Era*, 2001 U. ILL. L. REV. 173, 189 (citing Ken Howard, *The Genomics Gold Rush*, 283 SCI. AM. 59 (2000)).

17. *See generally* Eileen M. Kane, *Splitting the Gene: DNA Patents and the Genetic Code*, 71 TENN. L. REV. 707, 710 & n.13 (2004) (citing David J. Galas, *Making Sense of the Sequence*, 291 SCIENCE 1257 (2001), for the proposition that there are approximately 30,000 human genes, which code for over 90,000 proteins, and that noncoding parts of the DNA may have regulatory and structural roles).

18. *Id.* at 724. Kane gives the example of CCR5, a gene sequence which was patented on the basis of its coding for a protein that functioned as a cell surface receptor. When variants of the protein were later found to confer resistance to HIV infections, the patentee was suddenly put in control of significant research into AIDS therapeutics. *Id.* at 720; *see also id.* at 722 (noting the many ways in which the one-gene, one-protein paradigm is not representative of the real complexity in genomic and proteomic science); *id.* at 723 (explaining how a single gene can produce a variety of products).

As the potential breadth of these patents became increasingly clear, the Federal Circuit set out to cure the problem of extraordinary coverage. Patent law prohibits patents on inventions that are not new, that are within the grasp of a person of ordinary skill in the art, or that are inadequately placed in the domain of knowledge.[19] The Federal Circuit thus addressed extraordinary coverage through its interpretation of the level of skill in the art, which is determinative of obviousness, enablement, and disclosure.[20] By assuming that the level was low, the Federal Circuit could have its cake and eat it too: It could ignore developments in the field that might have rendered specific gene and protein sequences obvious,[21] and at the same time, it could interpret the enablement and written description requirements to narrow patent scope (or, perhaps more precisely, to invalidate patents as a way to motivate patentees to draft claims narrowly).[22] In theory, this practice ought to have restored competition to the field, as infringement of gene patents could now be avoided by making small modifications to the gene. For example, companies seeking to manufacture pharmaceuticals could get around a patent by altering the protected nucleotide sequences in ways that generated the same (or a functionally similar) product.

A similar story is applicable to proteins. As with genes, patents on proteins are potentially quite broad because proteins have a variety of effects in the body and can often be manipulated in many different ways. With a narrow view of the level of skill in the art of proteomics, patents could be granted. However, applying that level of skill to the issues of enablement and written description potentially worked to narrow each patent considerably. For example, a scientist interested in examining the interaction between a drug candidate and a specific portion of a patented protein could alter a part of the protein that is distant from the point of interaction. As long as the newly altered protein does not fall within the narrow scope of the claim, the scientist could then do his or her research without infringing.

19. 35 U.S.C. §§ 101, 102, 103, 112 (2000 & Supp. 2005).

20. *See id.* §§ 103, 112.

21. The key cases on this issue are over ten years old and deal with research that occurred even earlier. *See In re* Deuel, 51 F.3d 1552 (Fed. Cir. 1995); *In re* Bell, 991 F.2d 781 (Fed. Cir. 1993).

22. *See* Regents of the Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559, 1568 (Fed. Cir. 1997) (holding that the written description of the nucleotide of rat insulin does not cover human insulin); *In re* Wright, 999 F.2d 1557 (Fed. Cir. 1993) (holding that the disclosure of one working example of a vaccine against one RNA virus does not enable the production of vaccines against all RNA viruses); Janice M. Mueller, *The Evolving Application of the Written Description Requirement to Biotechnological Inventions*, 13 BERKELEY TECH. L.J. 615 (1998); *see also* Patent & Trademark Office, Written Description Requirement, 66 Fed. Reg. 1099 (Jan. 5, 2001).

In practice, however, this approach created significant new problems. This form of narrowing is useless in certain applications, particularly in fundamental research, where studying an organism's *actual* genes and proteins is critical to understanding functionality. And significantly, patentees have not stood back idly, watching these kinds of substitutions erode the value of their patents in either research or industrial settings. First, they changed their strategies. When patenting genes, they claim, for example, nucleotide sequences "substantially" identical, or "X% identity" to the sequences they had discovered or all sequences that encoded the proteins they had found.[23] To capture a full array of proteins, patentees claim genes and then add a variety of reach-through claims, claiming the sequences with which the complementary strand could hybridize along with the proteins these sequences could produce.[24]

Second, patentees have successfully challenged the ambit of the common law experimental use defense, making it more difficult for potential competitors to do the work needed to find functionally equivalent products. Thus, in the last few years, the Federal Circuit has obliterated the distinction between using a patented invention at the research stage as opposed to the manufacturing stage, or between commercial and noncommercial purposes. As a result, there is no longer a general right to use a patented invention to design around it,[25] to demonstrate the effectiveness of the

23. For examples of patents utilizing these claiming strategies, see U.S. Patent No. 6,593,135 (filed May 19, 2000) and U.S. Patent No. 6,777,590 (filed Dec. 22, 1999).

24. For examples of patents utilizing this claiming strategy, see U.S. Patent No. 6,872,579 (filed July 9, 2003) and U.S. Patent No. 6,426,198 (filed June 1, 1999). It remains to be seen how successful these strategies are in terms of patent validity, but a recent empirical study suggests that the applicable scope of *Lilly* remains uncertain. *See* Christopher M. Holman, *UC v. Eli Lilly* After Eight Years: An Empirical Study of the Impact of the *Lilly* Doctrine on the Patenting of Biotechnology and Chemical Inventions 8 (forthcoming paper presented at the 5th Annual Intellectual Property Scholars Conference at Cardozo School of Law, Aug. 11–12, 2005), *available at* http://justinhughes.net/ipsc2005/papers/Paper-HOLMAN.doc (citing *In re* Wallach, 378 F.3d 1330 (Fed. Cir. 2004) (rejecting a claim to all sequences encoding a particular protein, but implicitly permitting a patent on a protein claimed in terms of its partial amino acid sequence, function, molecular weight, and purification parameters to issue)); *see also* Noelle v. Lederman, 355 F.3d 1343 (Fed. Cir. 2004) (upholding a claim to all monoclonal antibodies capable of recognizing a specified mouse antigen); Enzo Biochem, Inc. v. Gen-Probe Inc., 323 F.3d 956, 966 (Fed. Cir. 2002) (holding that the reference to the deposit of three sequences provides sufficient disclosure to support a claim to all molecules that bind to specified chromosomal DNA). Holman also describes decisions of the Board of Patent Appeals and Interferences upholding claims disclosed in terms of percent identity to specified reference sequences and various hybridization claims. *See, e.g.*, Capon v. Eshhar, 418 F.3d 1349, 1358 (Fed. Cir. 2005) ("When the prior art includes the nucleotide information, precedent does not set a *per se* rule that the information must be determined afresh.").

25. *See* Soitec, S.A. v. Silicon Genesis Corp., 81 F. App'x 734 (Fed. Cir. 2003) (nonprecedential).

design-around,[26] or, perhaps, even to study the scientific principles underlying it.[27] Judge Rader would go even further: "The Supreme Court's recent reiteration that infringement does not depend on the intent underlying the allegedly infringing conduct, to my eyes, precludes any further experimental use defense."[28]

The bottom line is a system that is becoming increasingly unworkable. The new claiming strategies have produced patents that are more expensive to draft and harder for the U.S. Patent and Trademark Office (PTO) to examine. Their boundaries are difficult to discern, making issued patents rife with opportunities for challenge. There is anecdotal evidence that a new form of defensive patenting is taking hold, what might be called "arms race patenting," where patents are sought out of fear that a firm without a strong patent portfolio will not have the ammunition it needs to settle the infringement suits its competitors may file.[29] These new defensive patents do nothing to advance science.[30] Instead, they create thickets of rights that are ever more costly to negotiate and license.

University researchers are particularly concerned about the impact of these developments on their research programs because university researchers do not typically engage in the type of marginal work that yields defensive patents. Nor do they choose projects for their commercial potential or enjoy the authority to monitor their colleagues to ensure that they are not engaged in infringing activity. However, the broader research community, where firms routinely analyze their researchers' activities to make sure that patents are not being violated, is also affected. As Judge Newman recently put it:

> Were . . . research [to improve upon the patent, find a new use for it, or to modify or design around it] subject to prohibition by the patentee the advancement of technology would stop, for the first

26. Embrex, Inc. v. Serv. Eng'g Corp., 216 F.3d 1343 (Fed. Cir. 2000).

27. Madey v. Duke Univ., 307 F.3d 1351, 1362 (Fed. Cir. 2002). The reach of this case is unclear because the patent at issue was being used as a research tool, not as an object of research. However, the court's broad wording, to the effect that universities engaged in their normal business of conducting research are not entitled to rely on a research exemption, is worrying. In *Merck KGaA v. Integra Lifesciences I, Ltd.*, 125 S. Ct. 2372 (2005), the Supreme Court interpreted the statutory exemption, 35 U.S.C. § 271(e)(1) (2000), broadly, but it affects only uses directly related to submitting information to the Food and Drug Administration (FDA). In the Court's words:

> Basic scientific research on a particular compound, performed without the intent to develop a particular drug or a reasonable belief that the compound will cause the sort of physiological effect the researcher intends to induce, is surely not "reasonably related to the development and submission of information" to the FDA.

Merck KGaA, 125 S. Ct. at 2382.

28. *Embrex, Inc.*, 216 F.3d at 1353 (Rader, J., concurring) (citing, among other cases, *Warner-Jenkinson Co. v. Hilton Davis Chemical Co.*, 520 U.S. 17, 34 (1997)).

29. *See* Wesley M. Cohen et al., *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* 28–31 (Nat'l Bureau of Econ. Research, Working Paper No. 7552, 2000), *available at* http://papers.nber.org/papers/W7552.v5.pdf.

30. *Id.* at 29–30.

> patentee in the field could bar not only patent-protected competition, but all research that might lead to such competition, as well as barring improvement or challenge or avoidance of patented technology. Today's accelerated technological advance is based in large part on knowledge of the details of patented inventions and how they are made and used. Prohibition of research into such knowledge cannot be squared with the framework of the patent law.[31]

Clearly, new thinking is required. If the problems are to be solved, it is necessary to reexamine the drug development pipeline to better understand which of its steps are straightforward and which are difficult, when patenting makes sense and where patenting matters.

II. THE REALITIES OF STRUCTURAL BIOLOGY AND GENOMICS AND THEIR ROLES IN DRUG DISCOVERY

As noted earlier, changes in biotechnology are vast; for expository reasons, we restrict our account to structural biology and structural genomics, two fields that are particularly important in rational drug development. Structural biology is the study of biological processes through an analysis of the molecular structure of proteins. That is, a structural biologist begins by identifying specific protein molecules that she believes to be involved in a biological process. She will isolate these proteins, determine their three-dimensional structures through biophysical methods (such as X-ray crystallography and nuclear magnetic resonance spectroscopy), and use what she learns to better understand how these proteins function biologically. Structural genomics is the study of the structures of proteins on a genomic scale.[32] A structural genomicist will use bioinformatics techniques to analyze nucleotide sequences in the genome that can express[33] proteins whose three-dimensional structures are unknown. He will then use biological techniques to clone and express the gene, purify the resulting protein, determine its structure, and then utilize biochemical and biophysical methods to elucidate functionality. In a sense, these two approaches are complementary to each other. The former is hypothesis driven,

---

31. Integra Lifesciences I, Ltd. v. Merck KGaA, 331 F.3d 860, 875 (Fed. Cir. 2003) (Newman, J., concurring in part, dissenting in part), *vacated*, 125 S. Ct. 2372 (2005) (not reaching the common law research exemption issue addressed by Judge Newman).

32. Structural genomics is considered "the next step beyond the human genome project," the project that generated the map of the human genome. *See, e.g.*, Nat'l Ctr. for Biotechnology Info., A Gene Map of the Human Genome, http://www.ncbi.nlm.nih.gov/SCIENCE96/ (last visited Jan. 14, 2006).

33. Which is to say, that can actually produce the protein which the gene encodes.

in that the scientist starts with a theory about what is going on biologically and generates data. In the latter, the scientist looks for gaps in the existing data with the goal of ultimately formulating a theory.

As to both areas, the early cases were not entirely wrong. In the early period of modern biology, each of the steps from gene to function to structure to pharmaceutical was hard-fought. But as Part II.A demonstrates, both structural biologists and structural genomicists have advanced their techniques dramatically and now use "high throughput methods," which are procedures that have been developed to automate structural determinations. However, as Part II.B suggests, although both sets of scientists end up with protein structures that can then be used as targets for drug development, discovery of new drugs remains a challenging and expensive proposition.

A. Structural Biology and Genomics

Before the 1970s, both gene and protein sequencing were—as the early case law assumes—labor-intensive processes, involving chemically analyzing the content and order of the nucleotides in the gene and the amino acids in the protein. These methods were, indeed, so difficult that they were worthy of Nobel Prizes. For example, Fred Sanger won two, first for protein sequencing in 1958 and then for nucleic acid sequencing in 1980 (the year that *Chakrabarty* was decided).[34]

At that time, the next stage in development—learning how these proteins function, elucidating their three-dimensional architecture ("fold"), and measuring the "pockets" into which ligands (potential pharmaceuticals) can bind—was equally challenging.[35] In those early days, experimental determination of a three-dimensional structure began by purifying proteins from material extracted from native sources such as pigs' hearts or horses' blood. It next required crystallization of the resulting protein from many milligrams of material and with months of trial-and-error experiments. Data collection was done with relatively low-intensity X-ray sources; the data were collected on film and read either by eye or with an electronic scanner. The data were then analyzed using computers with far less power than the machines of today. From this data, protein models were constructed manually in physical space—often large physical spaces—using materials like brass or plastic.[36]

34. DONALD VOET & JUDITH G. VOET, BIOCHEMISTRY 825 (1990).

35. *See generally* T.L. BLUNDELL & L.N. JOHNSON, PROTEIN CRYSTALLOGRAPHY (1976).

36. For an example of an early protein model, see Ohlendorf Lab: Pictures From the Lab, http://biosci.cbs.umn.edu/bmbb/ohlen_lab/photoGallery/original/FredsFolly.jpg (last visited Oct. 31, 2005).

Much of this has changed. The labor-intensive methods of development have given way to technology-driven automated processes that make development a far easier process. Now, nucleic acid sequencing is routine and automated; protein sequences can be derived from gene sequences using computationally driven bioinformatics methods.[37] And although the goal of determining protein structure from sequences using purely in silico methods is as yet unrealized, high throughput experimental methods, which are themselves based on tremendous technological breakthroughs, make structural determinations far easier. Thus, recombinant DNA methodologies make it possible to obtain large quantities of pure protein by first cloning and expressing the gene and then purifying the resulting protein; crystallization is done using tiny amounts of material with computerized robots; extremely high-intensity X-ray sources, produced at specialized national centers,[38] generate data in a matter of hours; and highly sensitive image detectors collect the data, which are analyzed computationally. Using modern computational methods, it is possible to model an entire structure in a matter of days. Furthermore, these models can be fully visualized on computers.[39]

As a result of these advances, there are now large numbers of sequences and structures that are known. The quantity of knowledge that can be gleaned from data analysis has led scientists to organize international efforts to collect proteomic and genomic data in an orderly fashion. There are archival data banks of gene sequences such as GenBank,[40] DNA Data Bank of Japan (DDBJ),[41] and European Molecular Biology Laboratory Nucleotide Sequence Database (EMBL-Bank)[42] that together contain over fifty million DNA sequences.[43] SwissProt (now Uniprot) contains fully annotated protein sequences.[44]

37. Russ B. Altman & Jonathan M. Dugan, *Defining Bioinformatics and Structural Bioinformatics*, *in* STRUCTURAL BIOINFORMATICS 3 (Philip E. Bourne & Helge Weissig eds., 2003).

38. For a list of these so-called synchrotron beamlines, see BioSync, U.S. Synchroton Beamlines for Structural Biology, http://biosync.rcsb.org/ (last visited Oct. 31, 2005).

39. Paul D. Adams et al., *Computational Aspects of High-Throughput Crystallographic Macromolecular Structure Determination*, *in* STRUCTURAL BIOINFORMATICS, *supra* note 37, at 75.

40. Nat'l Ctr. for Biotechnology Info., GenBank Overview, http://www.ncbi.nlm.nih.gov/Genbank (last revised Nov. 7, 2005).

41. DNA Data Bank of Japan, What Is DDBJ, http://www.ddbj.nig.ac.jp (last visited Oct. 31, 2005).

42. Euro. Bioinformatics Inst., EMBL Nucleotide Sequence Database, http://www.ebi.ac.uk/embl/index.html (last visited Oct. 31, 2005).

43. *See, e.g.*, Nat'l Ctr. for Biotechnology Info., *supra* note 40.

44. UniProt Consortium, UniProt: The Universal Protein Resource, http://www.pir.uniprot.org/ (last visited Jan. 23, 2006).

The Protein Data Bank (PDB), which started in 1971[45] with just seven protein structures, now contains the structures of over 30,000 proteins.[46]

There are also thousands of databases derived from these archival databases. For example, there are databases that classify the elements of protein structure based on common features, such as folds.[47] There are also data resources that relate sequence and structure to function. Protocols for representing and validating these data have also been developed. These databases have become so important that funding agencies and journals require submission of data to them.[48] While submission has thus become the norm in academia, there are mechanisms in the private sector that likewise encourage submission. Firms engaged in doing structures for commercial purposes have also formed consortia (arguably, to preempt patenting) to put single nucleotide polymorphisms (SNPs)—alterations of genetic sequences—into the public domain.[49]

---

45. *Crystallography: Protein Data Bank*, 233 NATURE NEW BIOLOGY 223 (1971).

46. New initiatives are continuously proposed. For example, Francis Collins recently suggested the creation of a database of the gene sequences of malignant tumors, with the idea that it would allow scientists to find the mutations associated with various cancers and single them out for drug development. *See* Andrew Pollack, *Human Genome Project Is Proposed to Fight Cancer*, N.Y. TIMES, Mar. 28, 2005, at A11.

47. Helge Weissig & Philip E. Bourne, *Other Structure-Based Databases*, *in* STRUCTURAL BIOINFORMATICS, *supra* note 37, at 217.

48. For sample journal policies, see Philip Campbell, *New Policy for Structural Data*, 394 NATURE 105 (1998) and Nature Publ'g Group, *Rules of Genome Access*, 404 NATURE 317 (2000) ("*Nature*'s policy is that human sequence data should be deposited in a reliable, publicly available, unrestricted and free database."), *available at* http://www.nature.com/nature/archive/index.html; Am. Ass'n for the Advancement of Sci., General Policies for Authors: Conditions of Acceptance, http://www.sciencemag.org/about/authors/prep/gen_info.dtl#conditions (last visited Jan. 23, 2006) ("Before publication, large data sets, including protein or DNA sequences, microarray data, and crystallographic coordinates, must be deposited in an approved database and an accession number provided for inclusion in the published paper . . . ."). Several journals follow the policy of the International Union of Crystallography. *See* Int'l Union of Crystallography, *Guidelines for the Deposition and Release of Macromolecular Coordinate and Experimental Data*, D56 BIOLOGICAL CRYSTALLOGRAPHY 2 (2000), *available at* http://journals.iucr.org/d/issues/2000/01/00/issconts.html. For links to other journal policies, see Research Collaboratory for Structural Bioinformatics (RCSB), PDB Data Release Policies, http://deposit.pdb.org/depoinfo/PDB_deposition_release_policies.html (last visited Jan. 14, 2006). For sample funding agency policies, see RCSB, NIH Policy Relating to Deposition of Atomic Coordinates Into Structural Databases, http://deposit.rcsb.org/depoinfo/deposition_release_policies.html (last visited Jan. 14, 2006) ("[A]tomic coordinates should be deposited for immediate release upon publication of [a sponsored] research article."); Nat'l Human Genome Research Inst., NHGRI Policy for Release and Database Deposition of Sequence Data (Dec. 21, 2000), http://www.genome.gov/10000910.

49. Single nucleotide polymorphisms (SNPs) are genetic variations that can account for different pharmacological responses among individual patients. Accordingly, they are extremely important to drug development. *See* SNP Consortium Ltd., Single Nucleotide Polymorphisms for Biomedical Research, http://snp.cshl.org/ (last updated Aug. 10, 2004) (noting the firms taking part in the consortium); *see also* Rebecca S. Eisenberg, *The Promise and Perils of Strategic Publication to Create Prior Art: A Response to Professor Parchomovsky*, 98 MICH. L. REV. 2358, 2364–65 (2000) (describing Merck's sponsorship of the Merck Genome Initiative, which created a public database of expressed sequence tags (ESTs), sequences of gene fragments that express proteins).

And the same is happening in structural genomics with the formation of a commercial consortium funded by Wellcome Trust.[50]

These data banks represent important advances for genomicists because they provide the core data for their bioinformatics analyses. Genomicists will check each of these data banks to determine the uniqueness of the gene sequences and proteins that they are examining. Structural biologists will also utilize these data banks to learn as much as they can about the systems that they are studying. For example, one result of extensive sequence and structure analyses of these data is that humans are now known to share their genetic endowment with other living species. Understanding what is happening at the molecular level in a monkey, for example, can be highly informative of what happens in a human—and vice versa. The shape of particular proteins, such as hemoglobin, is even more conserved in evolution than are the sequences. Thus, it is possible to develop a fairly accurate sense of the three-dimensional shape of a protein in one organism by knowing the shape of the same protein in another.

Now that so much is known about the sequences of nucleotides, the sequences of proteins, the shape of proteins, and the relationships among proteins in different species, it is tempting to assume—as was described at the outset—that drug development should now be easy. However, this knowledge is only one component of the process. In fact, the last decade has shown that rapid drug discovery will not be quickly or easily achieved—not because the first steps in the process are unimportant, but rather because the later steps are far from obvious.

B. Drug Discovery[51]

The identification of pharmaceutically active compounds using the fruits of structural biology continues to focus on the architecture of proteins, including their folds and the shapes of their pockets. In today's world, however, this usually involves a variety of automated and computational techniques for screening compounds that are potentially bioactive. First one hones in on a set of potential drug candidates, and then one investigates these candidates to determine their suitability as therapeutic agents.

Researchers will generally begin with libraries of small molecules that are likely to interact with a particular protein target, either because they are

50. *See* Wellcome Trust, Structural Genomics Consortium, http://www.wellcome.ac.uk/doc_WTD003502.html (last visited Oct. 31, 2005).

51. Melissa Feeney Wasserman contributed the following discussion, for which we are very grateful.

structurally similar to molecules of known bioactivity or because computational studies indicate that they are likely to have a shape that will fit a particular pocket of the target.[52] These molecules are then screened, either with automated high throughput screening techniques, or by examining X-ray crystal structures and comparing in detail the relationship between the architecture of the pocket and of the small molecule.[53] Compounds identified as active in one of these ways are then retested to confirm activity and selectivity.[54] Although the first screen may be relatively quick (since it is usually automated), the secondary evaluation—the goal of which is to select so-called "lead compounds" for further development—is time-consuming. It has been estimated that for every 5000 to 10,000 compounds screened there are about 250 lead candidates placed into preclinical testing.[55]

Next, lead compounds are "optimized": Their structures are altered to create the most efficacious candidates to use for drug development.[56] Lead optimization is usually an iterative process involving ligand synthesis and compound evaluation. Compound evaluation can include computational techniques, including both computer modeling and laboratory experiments. Two computational methods frequently used in lead optimization are rational structural design[57] and pharmacophore analysis.[58] In the former, the constituent side chains or functional groups of the lead compounds are varied to provide a series of virtual compounds to be modeled. The interactions of these compounds and the binding site of the protein are examined computationally to estimate how well the molecule will fit the receptor (or active) site of the protein target. Scoring systems are set up to calculate quantitatively how well the ligand docks with the active site in order to identify promising

52. K.H. Bleicher, *Chemogenomics: Bridging a Drug Discovery Gap*, 9 CURRENT MEDICINAL CHEMISTRY 2077 (2002).

53. Irwin D. Kuntz, *Structure-Based Strategies for Drug Design and Discovery*, 257 SCIENCE 1078 (1992); Suzanne B. Shuker et al., *Discovery High-Affinity Ligands for Proteins: SAR by NMR*, 274 SCIENCE 1531 (1996).

54. A. Alanine et al., *Lead Generation—Enhancing the Success of Drug Discovery by Investing in the Hit to Lead Process*, 6 COMBINATORIAL CHEMISTRY & HIGH THROUGHPUT SCREENING 51, 56–58 (2003).

55. Joseph A. DiMasi et al., *Cost of Innovation in the Pharmaceutical Industry*, 10 J. HEALTH ECON. 107 (1991).

56. A.S. Verkman, *Drug Discovery in Academia*, 286 AM. J. PHYSIOLOGY—CELL PHYSIOLOGY C465, C469 (2004).

57. *See* Paul D. Lyne, *Structure-Based Virtual Screening: An Overview*, 7 DRUG DISCOVERY TODAY 1047 (2002).

58. *See* K.-J. Schleifer & E. Tot, *Pharmacophore Modelling of Structurally Unusual Diltiazem Mimics at L-Type Calcium Channels*, 14 J. COMPUTER AIDED MOLECULAR DESIGN 427 (2000).

drug candidates.[59] Most algorithms utilized consider both structural and functional interactions, such as steric fit, hydrogen bonding, and hydrophobic interactions.

Similar to the rational structural design method, the pharmacophore analysis aims to predict whether particular small molecules will bind to the target protein. In the pharmacophore method, the binding site of the protein is reduced to a minimal unit that is responsible for the bioactivity. It will usually consist of a combination of hydrogen bond donors/acceptors, hydrophobic groups, and other functional groups.[60] This minimal unit of functional groups is defined in three-dimensional space and constitutes the pharmacophore. The pharmacophore is then used to examine the allowed placement of groups in structural analogs of the lead compounds. Compounds whose functional groups have the right geometry are identified as promising candidates.

Promising compounds are then synthesized and evaluated experimentally in the laboratory. After synthesis, binding assays, co-crystallization of the compound with the protein, and X-ray structural studies are performed. These studies use the modern rapid methods described above. The compound's pharmacological properties, such as absorption, distribution, metabolism, excretion, and toxicity are also assessed. In the later stages of lead optimization, the focus of the experimental evaluation shifts towards in vivo testing in disease-related animal models. At this time, compounds undergo extensive testing for possible side effects, such as mutagenic, teratogenic, or cardiovascular effects, together with rigorous testing of stability and metabolism.

Lead optimization almost always involves the efforts of a team of researchers who work together to process results from compound evaluation studies with a view to establish a structure-function relationship. Knowledge of protein-ligand structure and ligand pharmacological properties are utilized by researchers to refine the structure of lead compounds with the aim to improve target-binding. The refined lead compound is then further refined in a reiterative optimization process.

Lead optimization is cost intensive, typically costing anywhere from two to four million dollars,[61] and can take several years to complete.[62] Significantly, decisions made during the process are decidedly nontrivial. The research team must weigh the various factors appropriately to ensure that desirable properties

59. Jonathan D. Hirst, *Predicting Ligand Binding Energies*, 1 CURRENT OPINION DRUG DISCOVERY & DEV. 28, 29–30 (1998).

60. Schleifer & Tot, *supra* note 58, at 427–28.

61. Verkman, *supra* note 56, at C466.

62. Hans-Jürgen Federsel, *Logistics of Process R & D: Transforming Laboratory Methods to Manufacturing Scale*, 2 NATURE REV. DRUG DISCOVERY 654, 657 (2003).

are maximized while essential properties are maintained. Thus, the multi-dimensional optimization problem requires a research team made up of diverse disciplines who must work together to determine which candidates should be pursued and which should be discarded. Eventually this team will decide which compounds, if any, will become drug candidates.

Even when lead optimization is completed, there is no guarantee that the candidate drug will survive the required and extensive testing in humans. In fact, only one in ten candidate drugs prove successful in all three phases of clinical trials.[63] After all, it is not only important that the compound exhibit the biochemical function sought; the drug must also be efficacious in humans and not harm the patient in unanticipated ways.

In sum, knowledge of gene and protein sequences, and even three-dimensional structures, is not the end of the drug development process. Nor is knowledge of the structure of a protein-ligand complex. While this means that drug development is no longer a matter of brute trial and error, it is still far upstream in the process of finding new therapies. Indeed, despite many efforts to take a rational, structure-based approach to drug design, to date there are only a few examples of drugs that have actually been brought to market using these methods:[64] human immunodeficiency virus (HIV) drugs such as amprenavir (Agenerase) and neflfinavir (Viracept), which were developed using crystal structure of HIV protease;[65] the flu drug zanamivir (Relenza), which involved extensive modeling based on the crystal structure of neuraminidase;[66] and resistance problems to the leukemia drug imatinib (Gleevec) have been better understood by analyzing the crystal structure of c-ABL.[67] It is no wonder, then, that a more realistic understanding of the limits of these advances has brought many changes on the business side of biotechnology research.

There are likely some firms that continue to believe they can earn profits by developing high throughput sequencing and structure determination capabilities, utilizing these capabilities to find new proteins, and then using the proteins to develop drugs or license the proteins out for use as research targets. However, it has become increasingly clear that simply knowing sequences

63. Tony Kennedy, *Managing the Drug Discovery/Development Interface*, 2 DRUG DISCOVERY TODAY 436, 442 (1997).

64. Tudor I. Oprea, *Chemical Space Navigation in Lead Discovery*, 6 CURRENT OPINION CHEMICAL BIOLOGY 384, 384 (2002).

65. Jonathan Greer et al., *Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design*, 37 J. MEDICINAL CHEMISTRY 1035, 1036 (1994).

66. Joseph N. Varghese, *Development of Neuraminidase Inhibitors as Anti-influenza Virus Drugs*, 46 DRUG DEV. RES. 176, 185–86 (1999).

67. Thomas Schindler et al., *Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase*, 289 SCIENCE 1938 (2000).

and even structures is not sufficient to guarantee a source of economic return, either from the discovery of a new therapeutic or from diagnostics. Accordingly, firms have migrated away from that business model, leading to a diversity of approaches for exploiting biomedical information.[68] Some firms specialize in finding new ways to screen drugs and then exploit these techniques to find lead candidates and, ultimately, new drugs. Other firms realize that they lack the expertise and resources to effectively mine and license the research opportunities that protein structures represent.[69] These firms rely on shifting alliances to create the capacity needed to turn genetic and proteomic knowledge into commercial products.[70] Other businesses have chosen to focus on manufacturing "value added" technologies. These include reagents that facilitate cloning and manipulating genetic and proteomic materials across a variety of environments; screening equipment, such as microchips embedded with an array of genome or proteomic material; animal models; analytical computer software; and laboratory equipment.

The resulting proliferation of patents in a context of disparate business plans, intellectual goals, and business and legal sophistication, has given rise

---

68. An example of this drift can be seen in the way that these companies describe their businesses in their securities filings. For example, Myriad Genetics described its business plan in 1997:

> Myriad Genetics, Inc. ("Myriad" or the "Company") is a leader in the discovery and sequencing of genes related to major common diseases. The Company utilizes analyses of extensive family histories and genetic material, as well as a number of proprietary technologies, to identify inherited gene mutations which increase the risk to individuals of developing these diseases.

Myriad Genetics, Inc., Annual Report (Form 10-K) 2 (filed Sept. 29, 1997), *available at* http://www.myriad.com/investors/secFilings.php (last visited Oct. 31, 2005). By 2004, the business was described as follows: "We are a leading biopharmaceutical company focused on the development and marketing of novel therapeutic and molecular diagnostic products. . . . We use this [genetic] information to guide the development of new healthcare products that treat major diseases and assess a person's risk of disease later in life." Myriad Genetics, Inc. Annual Report (Form 10-K) 3 (filed Sept. 10, 2004), *available at* http://www.myriad.com/investors/secFilings.php (last visited Oct. 31, 2005). Another example is Celera, which has moved from a sequencing company into drug development. *See* Maureen McDonough, *Celera Releases Genome Data at Last*, BIO-IT WORLD, June 8, 2005, http://www.bio-itworld.com/newsitems/2005/06-05/06-09-05-news-celera (reporting on Celera's release of its genome database). For Celera's current mission statement, see Celera, About Us, http://www.celera.com/celera/about (last visited Oct. 31, 2005).

69. *See* Avital Bar-Shalom & Robert Cook-Deegan, *Patents and Innovation in Cancer Therapeutics: Lessons From CellPro*, 80 MILBANK Q. 637 (2002).

70. Arti K. Rai, *Fostering Cumulative Innovation in the Biopharmaceutical Industry: The Role of Patents and Antitrust*, 16 BERKELEY TECH. L.J. 813, 816–18 (2001); *see also* Walter W. Powell, *Networks of Learning in Biotechnology: Opportunities and Constraints Associated With Relational Contracting in a Knowledge-Intensive Field*, *in* EXPANDING THE BOUNDARIES OF INTELLECTUAL PROPERTY: INNOVATION POLICY FOR THE KNOWLEDGE SOCIETY 251, 258 (Rochelle Dreyfuss et al. eds., 2001) [hereinafter EXPANDING THE BOUNDARIES]; Walter W. Powell, *Inter-organizational Collaboration in the Biotechnology Industry*, 152 J. INSTITUTIONAL & THEORETICAL ECON. 197, 205 (1996).

to significant licensing difficulties.[71] In many cases, negotiations are slow relative to the rate at which scientific research proceeds. Attempts to streamline the process with standardized legal forms have generally failed, perhaps because laboratory scientists do not want to take the time to understand them or because lawyers find each deal too unique to make standardization appropriate. And because much of the work in these fields has occurred in university settings and has reached fruition in the post–Bayh-Dole era of university involvement in patenting,[72] many licensing negotiations have occurred under the auspices of nascent—and in some cases, poorly funded—technology transfer offices, and have not resulted in significant returns to universities in most cases.[73] Their patents have also engendered conflicts of interest; created tensions among researchers, faculty, students, and university administrations;[74] and given courts hearing patent disputes reason to equate universities with commercial players, often to the detriment of important institutional values.[75]

III. PATENT LAW

The previous discussion highlights the reasons why there is a need to reconsider patent law as it applies to structural biology and genomics. It is now clear that the one gene/one protein/one disease paradigm is simplistic and that patents on genes and proteins can cover a broad array of research opportunities. These patents can be significantly different from rights in other fields because they have no substitutes, at least not for certain applications. Furthermore, these innovations are many steps removed from commercial products. The experiments that must be conducted to find marketable drugs are risky, involve multiple sorts of expertise, and are difficult enough to be considered inventive—and therefore patentable—in their own right. Thus, the question for innovation policy is where in the pipeline patenting actually belongs.

71. Rebecca S. Eisenberg, *Bargaining Over the Transfer of Proprietary Research Tools: Is This Market Failing or Emerging?*, *in* EXPANDING THE BOUNDARIES, *supra* note 70, at 223.

72. 35 U.S.C. §§ 200–212 (2000 & Supp. 2003) (allowing universities to retain rights to federally funded inventions).

73. Rebecca S. Eisenberg, *Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research*, 82 VA. L. REV. 1663, 172–14 (1996).

74. *See* Rochelle Cooper Dreyfuss, *Collaborative Research: Conflicts on Authorship, Ownership, and Accountability*, 53 VAND. L. REV. 1161 (2000) (noting, among other issues, the contentious issues regarding credit for work and access to tangible products and data); Lawrence M. Sung, *Navigating Uncharted Waters: Intellectual Property Rights Surrounding Genomics Research & Development Information*, 6 J. HEALTH CARE L. & POL'Y 194, 201 (2003).

75. Madey v. Duke Univ., 307 F.3d 1351 (Fed. Cir. 2002); *see also* Univ. of Rochester v. G.D. Searle & Co., 358 F.3d 916 (Fed. Cir. 2004); Griffith v. Kanamaru, 816 F.2d 624 (Fed. Cir. 1987) (rejecting the notion that a university's role as a research or teaching institution deserves special recognition).

For lawyers, the issue is how to interpret patent law to achieve that result. This part considers strategies for bringing the subject matter-inventiveness-scope-exemption trajectory into better alignment.

A. Patentable Subject Matter

Of all the issues in biotech patenting, the question whether genetic material and proteins are subject matter appropriate for patenting arguably deserves the most systematic attention. Unlike other issues that arise in patent litigation, the status of gene and protein discoveries as statutory subject matter has managed to escape review at all adjudicatory levels. Following *Chakrabarty*, the PTO refrained from denying patents on biotechnological inventions on subject matter grounds. And because those who are drawn into disputes over gene and protein patents tend to have vested interests in their availability, there is little incentive to challenge the statutory status of proteins and genes in district courts or the Federal Circuit.[76]

Nonetheless, a strong argument can be made that raw information about biological endowments should not be considered patentable unless the advance has end-product functionality (for example, as a therapy); that so long as such information is basically research information, it belongs in the public domain. After all, it is black letter intellectual property law that facts, principles of nature, and products of nature are not protectable. As the Supreme Court put it in *Funk Brothers Seed Co. v. Kalo Inoculant Co.*,[77] products of nature should be "free to all men and reserved exclusively to none."[78]

At first blush, it may seem curious that genes and proteins were ever considered subject to patent law. After all, they are significantly different from the invention discussed in *Chakrabarty*. The microorganism at issue in that case was a man-made product (created with gene splicing technology) with commercial value. However, simply denying patents to genes and

76. Kane, *supra* note 17, at 726. This view may, however, be changing. The Federal Circuit's decision to uphold the patentability of a method for detecting vitamin B deficiency by looking at the levels of certain amino acids, has been challenged in the Supreme Court on subject matter grounds. *See* Metabolite Labs., Inc. v. Lab. Corp. of Am. Holdings, 370 F.3d 1354 (Fed. Cir. 2004), *cert. granted in part*, 126 S. Ct. 601 (2005).

77. 333 U.S. 127 (1948).

78. *Id.* at 130; *see also* Feist Pub'ns, Inc. v. Rural Telephone Serv. Co., 499 U.S. 340, 344–45, 348 (1991) ("The most fundamental axiom of copyright law is that '[n]o author may copyright his ideas or the facts he narrates.'" (quoting Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 556 (1985))); *id.* at 348 ("[F]acts—scientific, historical, biographical, and news of the day . . . 'may not be copyrighted and are part of the public domain available to every person.'" (quoting Miller v. Universal City Studios, Inc., 650 F.2d 1365, 1369 (5th Cir. 1981))); O'Reilly v. Morse, 56 U.S. (15 How.) 62, 113 (1854).

proteins is not as straightforward as an analysis based on *Kalo* might suggest, for the material that is patented is not, as a technical matter, a product of nature. Rather, before patenting, genes and proteins are characterized by sequencing (for genes) and structural determination (for proteins). Both involve activities that have traditionally required isolation and crystallization. Courts have long drawn a line between genuine products of nature (not patentable) and products that have been enhanced through human intervention, typically through "processes of extraction, concentration, and purification of natural materials" (patentable).[79] Since isolation is a form of extraction, and crystallization is a form of concentration and purification, genetic material and proteins have been deemed patent-eligible.

In cases outside proteomics and genomics, the "human intervention" doctrine makes sense from both a pragmatic and an economic perspective. Isolated products are different in kind than products in nature because they have uses that the substances in nature lack. Furthermore, the efforts needed to identify such products and purify them enough to impart social value (as, for example, a medicine) represent the kind of investment that patent law was designed to encourage. But while this justification may have appeared appropriate in the early days of biotech research, when the expectation was that there was a tight, straightforward connection between a gene, a protein, and a specific therapy, it is unpersuasive in light of current understandings.[80]

For genes, the disjuncture between the justification for the human intervention doctrine and its application lies in the fact that value derives not from isolating genes, but rather from learning their informational content. Genes are, in other words, more like algorithms than like molecules. They are isolated as a step toward studying the instructions they give to the organisms in which they are found. Of course, proteins are molecules and function through chemical reaction, just as any molecule does. Furthermore, structural biologists purify proteins and crystallize them, and then utilize the crystalline form to determine their structure through X-rays and other methods. As we saw, however,

79. Merck & Co. v. Olin Mathieson Chem. Corp., 253 F.2d 156, 163 (4th Cir. 1958) (upholding the patent on vitamin B_{12}); *see also* Parke-Davis & Co. v. H.K. Mulford Co., 189 F. 95 (S.D.N.Y. 1911) (Hand, J.), *aff'd in part, rev'd in part*, 196 F. 496 (2d Cir. 1912) (upholding the patent on adrenalin).

80. Marty Adelman puts this another way. In his view, early cases such as *Amgen, Inc. v. Chugai Pharmaceutical Co.*, 927 F.2d 1200 (Fed. Cir. 1991), offered courts a choice between awarding the patentee a patent only on his true invention, a method for sequencing DNA, or a patent on the DNA itself. The first was often worth nothing because sequencing methods were changing rapidly; although the latter was worth far more than the patentee had contributed, it may have seemed the preferable alternative, given the value of the patentee's advance to human health. *See* 3 MARTIN J. ADELMAN, PATENT LAW PERSPECTIVES § 3.2[1] n.140 (2005).

crystallization of a protein is only a vehicle—a method for determining the protein's physical architecture when it *is* found in nature. That is, this work is undertaken precisely because the underlying assumption of structural biology, supported by substantial scientific evidence, is that the protein holds its shape. Once the geometric structure is elucidated by studying the crystal, it can be safely assumed that this is the shape found in nature.

For law, this suggests two related points. First, it is incorrect to think that a patent on a gene or protein would cover material that is different from what it is found in nature. For genes, the *information* is identical whether the gene is isolated or not; for proteins, the *shape* in a crystal is no different from the shape in nature. There is, in short, no difference "in kind," as in the usual human intervention case. Second, the value in finding sequences and structures is an intermediate value. It lies in learning about the gene or the protein in its natural context; it does not derive from isolation and purification as it does in the classic human intervention cases.[81] Rather, for genetic material and proteins, the effort that patent law is intended to encourage resides in the next set of steps that must be undertaken—converting that knowledge through long, intricate, and risky experimentation into commercial products.[82]

This is not to say that this material could never be patented. If a protein does have value by virtue of being isolated and turned into something different in kind from what is found in nature, then it should certainly be eligible for patent protection. For example, a protein that functions as a medicine, or a pure protein that is embedded into a chip and used for screening drugs, should be eligible for protection in the same way that any inventive pharmaceutical or research device is entitled to protection.[83] However, this does mean that it was a mistake to allow businesses to be built around patent protection of genetic material or proteins *as such*. Because genes and proteins are valuable by virtue of their function in nature, these patents can dominate broad swaths of far-upstream research opportunities and do exactly what *Kalo* and other subject matter cases sought to avoid: They endanger the efforts of "future

81. For example, the vitamin B_{12} compositions at issue in *Merck*, 253 F.2d at 165 n.6, were found to be superior to the liver extracts formerly used to treat permicious anemia. They were in a more abundant supply, were cheap, had a potency and dosage that could be precisely controlled, and were free of toxic substances.

82. In effect, the law has been turned on its head. In the usual case, characterization is an intermediate value: It merely demonstrates that a socially valuable task—purifying a natural compound—has been accomplished. For proteins, characterization has become an end in itself. By the mere fact of being characterized, the protein is regarded as eligible for protection.

83. These patents could well be broad enough to stifle fundamental research, but that problem can be dealt with in other ways, such as by limiting scope with a research exemption or compulsory license.

inventor[s], in the onward march of science," they "shut[ ] the door against inventions of other persons," and they allow "the patentee . . . to avail himself of new discoveries in the properties and powers of [inventions] which scientific men might bring to light."[84]

Unfortunately, much as many scientists would welcome a proposal to exclude genomic and proteomic inventions from patenting on theoretical as well as practical grounds, and much as laymen might see genes and proteins as "discoveries" rather than "inventions," this approach should probably be regarded as a horse well out of its barn.[85] Patents on existing genomic and proteomic information are unlikely to be withdrawn on a newly devised theory of patent law because to do so would unsettle investment-backed expectations. Limiting the exclusion to future advances would be a more viable approach, but imposing new and truncated definitions of patentable subject matter on existing technologies could jeopardize the capacity of the patent system to attract investments in future technologies. Excluding genes and proteins from patenting may also violate U.S. obligations under the TRIPS Agreement.[86] Besides, many of the suggestions made below (including the suggestions on utility and, particularly, on nonobviousness), should make these patents hard to acquire for other reasons. Accordingly, redefining statutory subject matter may not be necessary to free genomic and proteomic materials into the public domain.

B. Utility

Both the Constitution and the current patent law require patentees to disclose an end-use utility.[87] As Rebecca Eisenberg and Robert Merges suggest,

84. O'Reilly v. Morse, 56 U.S. (15 How.) 62, 113 (1854); *see also* NUFFIELD COUNCIL ON BIOETHICS, THE ETHICS OF PATENTING DNA 13 (2002), *available at* http://www.nuffieldbioethics.org/fileLibrary/pdf/theethicsofpatentingdna.pdf (last visited Oct. 31, 2005) (suggesting that the returns from these patents far exceed the technical contribution made by the patentees).

85. *See, e.g.*, *Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility*, 1300 OFF. GAZ. PAT. OFFICE 142 (Nov. 22, 2005), *available at* www.uspto.gov/web/offices/pac/dapp/opla/preognotice/guidelines101_20051026.pdf (making clear that an invention that "transforms an article or physical object to a different state or thing" or that "otherwise produces a useful, concrete, and tangible result," is patentable subject matter).

86. General Agreement on Tariffs and Trade—Multilateral Trade Negotiations (The Uruguay Round): Agreement on Trade-Related Aspects of Intellectual Property Rights, Including Trade in Counterfeit Goods art. 27(1), Dec. 15, 1993, 33 I.L.M. 81, 93–94 (1994) [hereinafter TRIPS Agreement] (prohibiting discrimination as to field of technology). *But see* Graeme B. Dinwoodie & Rochelle Cooper Dreyfuss, *International Intellectual Property Law and the Public Domain of Science*, 7 J. INT'L ECON. L. 431 (2004) (suggesting arguments why subject matter exclusions may not violate article 27.1).

87. U.S. CONST. art. I, § 8, cl. 8 (giving Congress authority to create patents for "useful Arts"); 35 U.S.C. § 102 (2000 & Supp. 2003); Brenner v. Manson, 383 U.S. 519 (1966).

this requirement "serves a timing function, leaving basic research discoveries in the public domain until they have yielded tangible benefits and have thereby left 'the realm of philosophy' and entered 'the world of commerce.'"[88] The PTO, which considered the issue of utility in the biotechnology context in 2001, attempted to operationalize the Eisenberg-Merges view. Its current guidelines emphasize the need to recite a "particular practical purpose" (in other words, the invention must have a "specific and substantial utility"), "credible by a person of ordinary skill in the art." The guidelines also specifically "exclude[ ] 'throw-away,' 'insubstantial,' or 'nonspecific' utilities."[89]

This is a high standard; it will surely exclude some raw biological information from patenting.[90] However, unless interpreted very narrowly, this standard will not exclude all structural genomic inventions from patenting because gene and protein sequences and structures will often be associated with specific characteristics of the organism in which they are found. For example, a gene associated with a particular disease (like sickle cell anemia) or one that confers a benefit (such as longevity or fertility) will—under current interpretations—have sufficient utility to meet the standard because they can be used as diagnostics. A protein that can be used to screen drugs might similarly qualify for protection.

Such patents are particularly likely to be problematic. Under current law, a patentee who identifies a single use for an invention obtains rights over all uses, including ones unknown at the time of patenting.[91] The rationale for this rule is not entirely evident. However, it is often argued that according broad rights is socially useful because it puts someone with proven technological sophistication in a position to orchestrate the development of the field to which the invention pertains.[92] But whatever the merits of that view

88. Rebecca S. Eisenberg & Robert P. Merges, *Opinion Letter as to the Patentability of Certain Inventions Associated With the Identification of Partial cDNA Sequences*, 23 AIPLA Q.J. 1, 6 (1995).

89. Utility Examination Guidelines, 66 Fed. Reg. 1092, 1098 (Jan. 5, 2001).

90. The Federal Circuit has affirmed a decision utilizing the guidelines. *See In re* Fisher, 421 F.3d 1365 (Fed. Cir. 2005). However, the court was careful to approve the *result*. While it agreed with the "PTO's standards for assessing whether a claimed invention has a specific and substantial utility," *id.* at 1372, the court did not, however, defer to the PTO's construction of the statute, *id.* (citations omitted) (noting that the guidelines "are not binding on this court, but may be given judicial notice to the extent they do not conflict with the statute"). Thus, the application of the guidelines as a general approach remains unclear.

91. *See, e.g.*, ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, PATENT LAW AND POLICY: CASES AND MATERIALS 237 (3d ed. 2002) (giving as an example a leather tanning agent later found to be effective as an anti-AIDS drug).

92. Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977). For an opposing view—that development is more effective if it is competitive—see Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990). Other justifications are discussed below, in connection with the discussion of scope.

generally, it is clearly inapplicable to those who do no more than find an association between a gene or protein and a disease. As suggested earlier, finding an association is relatively easy. Thus, the ability to recite such a utility is not a good indicator that the right holder also is capable of directing developments effectively. Indeed, it is not unknown for firms with broad patents to be unable or unwilling to mine or license them optimally.[93]

To prevent the acquisition of such unmanageable rights, the guidelines should therefore be interpreted so that claiming that structural information can function as a drug target or as a diagnostic—with no indication of what conditions are being targeted or diagnosed—is not considered a "particular practical purpose."[94] Moreover, we believe (as do patent offices in the United States, Europe, and Japan[95]) that experimental determination is critical. As was explained above, structures can be determined through comparative modeling techniques. But these can give scientists only an approximation of the overall shape of a protein; they are not accurate enough for purposes of drug development. Further experimentation is necessary before a structure can be used to learn about its ligand-binding properties.

As in connection with subject matter, some of the problems presented by these patents could be analyzed under other requirements as well. Thus, the problems caused when a single use confers power over all uses could be considered a question of patent scope and could arguably be better resolved by granting patents but limiting either their reach or the licensing freedom of their holders. By the same token, computer-mediated predictions may be better considered under the nonobviousness requirement. These issues are explored below.

93. *See supra* text accompanying notes 71 & 73. An example is the licensing practices of Myriad Pharmaceuticals in connection with its patent on breast cancer genes. *See* John P. Walsh et al., *Effects of Research Tool Patents and Licensing on Biomedical Innovation*, *in* PATENTS IN THE KNOWLEDGE BASED ECONOMY 285, 312 (Wesley M. Cohen & Stephen A. Merrill eds., 2003) (refusing to permit others to perform breast cancer tests using technology, even for research purposes).

94. Such a claim is no different from the "laundry list" of asserted utilities rejected in *Fisher*, 421 F.3d at 1377.

95. TRILATERAL PROJECT WM4, REPORT ON COMPARATIVE STUDY ON EXAMINATION PRACTICE RELATING TO SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) AND HAPLOTYPES (June 10–12, 2003), *available at* http://www.trilateral.net/projects/biotechnology/examination_snp/examination_on_snps_haplotypes.pdf; TRILATERAL PROJECT WM4, REPORT ON COMPARATIVE STUDY ON PROTEIN 3-DIMENSIONAL (3-D) STRUCTURE RELATED CLAIMS (Nov. 4–8, 2002), *available at* http://www.uspto.gov/web/tws/wm4/wm4_3d_report.htm (last visited Jan. 29, 2006).

C. Novelty and Nonobviousness

As noted above, patent law prohibits patents on inventions that are not new or that are within the grasp of a person of ordinary skill in the art.[96] These requirements protect the public domain from the withdrawal of information that is effectively already known. Equally important, they ensure that rights are awarded only in situations in which significant inventive effort has taken place.

At the time that patent law was first applied to biotechnology, it may have been reasonable to regard gene and protein sequences as meeting these requirements. Isolating and sequencing procedures were not fully developed. Furthermore, the code in which genes instruct the organism to generate proteins has many variations, making it difficult to relate known protein compositions to genetic sequences, and vice versa.[97] However, as we saw, structural biology has progressed significantly. To a large extent, sequencing is routine and automated; scientists can use the structure of a protein in one organism to accurately predict the structure of the same protein in another species.[98] Methods for elucidating the structure of small macromolecules is becoming routine. Given how rapidly the technology is progressing, it makes little sense to accord stare decisis effect to a finding on the level of skill in the art. Thus, even though sequencing and defining structures may once have been regarded as nonobvious in every case, sequences and structures that are now determined by high throughput processes, or by comparison to similar genes and molecules in other species, should no longer be considered inventive enough to qualify for patent protection.[99] As various trilateral conferences and studies of the National Research Council of the National Academies of

96. *See supra* note 19.

97. *See In re* Deuel, 51 F.3d 1552 (Fed. Cir. 1995); *In re* Bell, 991 F.2d 781 (Fed. Cir. 1993).

98. To be sure, the Federal Circuit has used the alleged difficulty of moving between species in its analysis of the enablement and written description requirement—that is, as a way to narrow the scope of genomic and proteomic patents. *See, e.g.*, Regents of the Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559 (Fed. Cir. 1997). While we agree that the scope of these patents should be narrow, the overview laid out above demonstrates why it is not particularly fruitful to narrow them by regarding minor variations in sequences as outside the scope of the claims. *See also* Dan L. Burk, *Biotechnology in the Federal Circuit: A Clockwork Lemon*, 46 ARIZ. L. REV. 441 (2004).

99. *In re* Crish, 393 F.3d 1253, 1254–55 (Fed. Cir. 2004), which affirmed the PTO's decision that a nucleotide sequence having promoter activity for the human involucrine gene (hINV) was anticipated by disclosure of a plasmid with the same hINV promoter activity, may signal a change in what the ordinary artisan is considered to understand.

Science (NAS) have suggested, this approach would have the added benefit of harmonizing U.S. law to the law of other industrialized countries.[100]

As we stressed earlier, changing the law will not eliminate patenting in the biotechnology field. But now that it is understood that sequencing is not the inventive step it was once thought to be, and that much work lies between structural determination and drug development, it is clear that there will be other patenting opportunities as researchers move along the pipeline. To be sure, some of that work will be, as the case law puts it, "obvious to try."[101] However, obvious to try is not a ground for refusing patent protection. The issue for patenting is how many alternatives are available and how likely each is to be successful.[102] Experience has demonstrated that drug development is still very much a hit-or-miss proposition. Thus, there will be other steps in the process that will qualify—and that are better suited—as the locus for patenting activity.

Admittedly, the task of continually updating skill levels is something that is extremely difficult for a court. This Article does not fully explore how to handle this situation; for these purposes, it is sufficient to note that a variety of effective approaches could be taken. The Federal Circuit could set out a methodology for lower courts to use.[103] In addition, the Federal Rules of Evidence give courts authority to appoint their own experts.[104] Although it can be hard for courts to find experts on their own, the various organizations active in the biotechnology field could create standing panels from which choices could be made. Another approach would be to defer to the PTO on this question.[105] Infusions of examiners skilled in the arts in which they work, supplemented by regular training sessions, give the PTO a unique perspective on what is known in the field. That knowledge could be augmented by periodic seminars with people active in the field or with standing expert panels.

---

100. NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., A PATENT SYSTEM FOR THE 21ST CENTURY 93 (Stephen A. Merrill et al. eds., 2004) [hereinafter NAS 2004 REPORT]; TRILATERAL PROJECT 24.1, BIOTECHNOLOGY COMPARATIVE STUDY ON BIOTECHNOLOGY PATENT PRACTICES COMPARATIVE STUDY REPORT (1998), *available at* http://www.trilateral.net/projects/biotechnology/patent_practices/biotechnology_patent_practices.pdf. For a comparison among patent offices, see Amanda S.Y. Lim & Andrew F. Christie, *Reach-Through Claims in Biotechnology: An Analysis of the Examination Practices of the United States, European and Japanese Patent Offices*, 3 INTELL. PROP. Q. 236 (2005).

101. *In re* O'Farrell, 853 F.2d 894, 903–04 (Fed. Cir. 1988).

102. *Id.*

103. Joseph P. Meara, *Just Who Is the Person Having Ordinary Skill in the Art? Patent Law's Mysterious Personage*, 77 WASH. L. REV. 267, 290–92 (2002).

104. FED. R. EVID. 706(a). *See generally* Howard M. Erichson, *Mass Tort Litigation and Inquisitorial Justice*, 87 GEO. L.J. 1983, 1986 (1999).

105. Arti Rai has written extensively on this approach. *See, e.g.*, Arti K. Rai, *Allocating Power Over Fact-Finding in the Patent System*, 19 BERKELEY TECH. L.J. 907 (2004); Arti K. Rai, *Engaging Facts and Policy: A Multi-institutional Approach to Patent System Reform*, 103 COLUM. L. REV. 1035 (2003).

As we saw, there are now international efforts to collect proteomic and genomic data in an orderly fashion; those who maintain these databases could serve as resources in this regard.

D. Disclosure

Patent law requires the patentee to describe the invention and the manner of making it clearly enough to enable a person skilled in the art to make and use it, and to understand exactly what is claimed to be within the patentee's exclusive authority.[106] The goals fostered by these requirements can be sorted—rather roughly—into two categories. The first, traditional, role of disclosure is informational in character. In this capacity, disclosure informs the PTO of the date on which the inventor was in true possession of the invention; it assures that the public will be able to utilize the invention after the patent expires; it helps disseminate technological information to the scientific community; and it delineates what the *patentee* regards as the metes and bounds of the invention.[107] In recent years, the Federal Circuit has, however, stressed a second function for disclosure. This is a definitional component: It serves as a measure of what the *law* regards as the appropriate scope of the patent claims.[108]

Common to both of those roles is the question of setting the level of skill in the art. If that level is too low, then specifications will be longer than needed to perform their informational function. More important, the definitional function will be undermined because the claims will be narrower than the actual teaching of the patent; in some cases, patents may be invalidated inappropriately. The upshot is that much of what was said in the nonobviousness section is applicable here: It is crucial to the proper operation of the system that the level of skill be regularly updated to take into account the rapid pace of development in these fields. Certainly, courts should not be bound by levels set in prior cases.[109]

106. 35 U.S.C. § 112 (2000).

107. Disclosure also requires the inventor to reveal the best mode of carrying out the invention; the NAS has recommended that this requirement be eliminated for reasons that are not relevant here. NAS 2004 REPORT, *supra* note 100, at 127.

108. *See, e.g.*, Regents of the Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559 (Fed. Cir. 1997); Cindy I. Liu, Gentry Gallery, Inc. v. Berkline Corp., 14 BERKELEY TECH. L.J. 123, 123 (1999).

109. It is encouraging to see that some judges have begun to articulate this problem expressly. For example, in *Capon v. Eshhar*, the court said:

> The law must be applied to each invention that enters the patent process, for each patented advance is novel in relation to the state of the science. Since the law is applied to each invention in view of the state of relevant knowledge, its application will vary with differences in the state of knowledge in the field and differences in the predictability of the science.

418 F.3d 1349, 1357 (Fed. Cir. 2005).

If this issue is reevaluated, then a further refinement, suggested by Rebecca Eisenberg, should also be considered.[110] In fact, the contexts in which skill in the art is significant differ from one another. For nonobviousness, the issue is whether other *researchers* would have discovered the advance based on existing knowledge. For disclosure, the question is whether *potential users* learn enough about the invention based on what is revealed in the patent. But there is no reason to assume that researchers and users have the same abilities to learn and utilize what they know. Arguably, scientists who enter research fields possess a degree of inventiveness and can extrapolate from what is known with a significant degree of effectiveness. In contrast, one might reasonably attribute a lower level of ingenuity to those who are in positions in which they are responsible only for implementing known inventions. Decoupling the inquiries and setting standards of skill geared to the circumstances in which the term is used could benefit the system. If not every extension of patented knowledge itself qualified for protection, then researchers would have more room to experiment.[111]

1. Informational Function

In addition to ensuring that the level of skill in the art is accurate, the patent must also be presented in a comprehensible manner. Disclosure can further informational goals only if it is in a format that people in the field find useful. As discussed, structural proteomic and genomic information has been collected into databases that are organized by professionals, with an eye toward making the information as transparent as possible.[112] Some patentees and lawyers voluntarily write patent applications that utilize the standards developed in connection with the relevant database.[113] However, because the representation of certain data, such as structural information, is highly complex, it would be preferable to read the requirement of enabling the "person skilled in the art" as *requiring* conformity.[114] For example, the

110. *See* Rebecca S. Eisenberg, *Obvious to Whom? Evaluating Inventions From the Perspective of PHOSITA*, 19 BERKELEY TECH. L.J. 885 (2004).

111. Graeme B. Dinwoodie & Rochelle Cooper Dreyfuss, *Patenting Science: Protecting the Domain of Accessible Knowledge*, *in* THE FUTURE OF THE PUBLIC DOMAIN IN INTELLECTUAL PROPERTY (L. Guibault & P.B. Hugenholtz eds., forthcoming 2006), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=698321.

112. *See supra* notes 40–46.

113. *See, e.g.*, U.S. Patent No. 6,826,488 (filed Mar. 23, 2000).

114. In the case where there is more than one database collecting similar information, the standards of any one of them could be considered sufficient.

PDB has clear specifications for how the data are to be formatted; the same specifications should be used in the patent application.[115]

Other forms of "informational piggybacking" are also possible. The PDB assigns a unique identification number to each structure deposited. Use of that identification in the patent would help the public because it would make it easer to determine the patent status of the protein more efficiently. Similarly, if deposit in the relevant data bank were made a requirement of the patent application process, the PTO could utilize the professional judgment of the data bank curators to determine whether the disclosure was accurate and informative to people in the field. For example, publication of a protein structure by the PDB is a strong indicator that the chemistry and geometry of the structure are consistent with protein science, and that the data are in a form that protein scientists understand.

To be sure, mandating involvement of data banks would require the cooperation of those in charge of maintaining them. In particular, inventions lose their patentability if they are disclosed for too long prior to the filing of the patent application.[116] Accordingly, the data bank would need to develop the capacity to delay publication of material that has been deposited. However, this problem would be easily solved, for some databases, such as the PDB, already have this capability. Various solutions to the publication problem could also be borrowed from other areas in which deposit is common.[117]

2. Definitional Function

Considerable controversy surrounds the Federal Circuit's decision to use the enablement and written description requirements to pressure patentees to draft claims of narrower scope.[118] It may be that relying on these provisions is

115. *See* RCSB, Protein Data Bank, RCSB PDB Data Validation and Deposition Services, http://pdb.rutgers.edu (last visited Oct. 31, 2005); RCSB, Protein Data Bank, PDB Documentation and Information, http://www.rcsb.org/pdb/info.html#File_Formats_and_Standards (last visited Oct. 31, 2005). In fact, the PTO is currently considering this issue. Patent & Trademark Office, Acceptance, Processing, Use and Dissemination of Chemical and Three-dimensional Biological Structural Data in Electronic Format, 70 Fed. Reg. 35,573 (June 21, 2005).

116. In the United States, this period is one year. 35 U.S.C. § 102(b) (2000). In other countries the grace period is shorter or nonexistent.

117. For example, seeds protected by the Plant Variety Protection Act must be deposited. 7 U.S.C. § 2422(4) (2000); *cf.* Enzo Biochem, Inc. v. Gen-Probe Inc., 323 F.3d 956 (Fed. Cir. 2002) (on significance of deposit). For a discussion of the timing of deposit and availability, see Elizabeth R. Hall & T. Ling Chwang, *Deposit Requirements for Biological Materials*, 14 HOUS. J. INT'L L. 565 (1992).

118. See, for example, the separate dissents issued in connection with the denial of rehearing in *University of Rochester v. G.D. Searle & Co.*, 375 F.3d 1303, 1314–24 (Fed. Cir. 2004) and the list of articles appended to Judge Rader's dissent, *id.* at 1314–24 (Rader, J., dissenting).

attractive because they depend on "skill in the art," which as a term of art, can be adjusted to achieve particular policy objectives regarding patent breadth. But while this approach may be appropriate to prevent patentees from gaining control over products that they have not in fact discovered (reach-through claims are examples), we have seen that it is inefficient (in that it does not guarantee access for those engaged in fundamental research) and creates several new problems.

First, because the commercial value of these products lies in their functionality, which for many purposes is undisturbed by trivial variations, tying the definition of the claims to a low level of skill in the art can deprive patentees of the full benefit of their technical contributions. Not only is the outcome unjust, it undermines the ability of the system to stimulate creativity. Every would-be inventor must discount the available rewards by the likelihood that a claim will be construed unfavorably, but if potential infringers do not think about costs in the same way, patents lose significant value without significantly increasing public access. Since people tend to be more risk-averse with regard to potential losses than potential gains,[119] it is particularly likely that access benefits will not offset motivational losses. Second, patentees try to avoid the erosion in value by engaging in various socially unhelpful efforts. As we have seen, they devise claiming strategies that impose costs, delays, and uncertainties on the system. Furthermore, they have persuaded the Federal Circuit to take a position on an experimental use defense that may be bad for science more generally.[120]

In sum, while we agree with the Federal Circuit's apparent intuition that genomic and proteomic patents are far too broad, the method it has chosen to narrow them needs to be reconsidered. In this regard, copyright law offers an important lesson. In copyright law, it is possible to make trivial variations in a work without disturbing its underlying functionality. Significantly, copyright regards such activity—paraphrasing—as infringing when it produces a product that potentially undermines the market for the original work.[121] So far, the Federal Circuit has rejected the notion that genomic inventions are information products that should be treated differently from other chemicals.[122] But perhaps the court would be willing to reconsider that position if other

119. Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCIENCE 453, 453–55 (1981).

120. *See* discussion *infra* Part III.E.2.

121. *Cf.* Nichols v. Universal Pictures Corp., 45 F.2d 119 (2d Cir. 1930) ("[T]he question is whether the part so taken is 'substantial,' and therefore not a 'fair use' of the copyrighted work; it is the same question as arises in the case of any other copyrighted work.").

122. *See, e.g.*, *Univ. of Rochester*, 375 F.3d at 1307 (Lourie, J., concurring).

methods for narrowing patent scope were available. Likewise, patentees may be more receptive to solving the social problems that genomic and proteomic patenting present if they were assured of gaining a full return from the technical contributions that they, in fact, made. Three ways to narrow patent scope are discussed below.

E. Scope

As noted earlier, traditional patent law gives patentees authority over all uses of the patented invention, including uses that were not disclosed in the specification. We previously suggested that one of the theories behind that approach—the view that patentees will orchestrate efficient development—appears inapplicable when patents can be awarded before substantial inventive work has been accomplished. Other justifications for broad patents—secrecy and the sufficiency of return on investment—may be equally inapposite when applied to biotechnology.

The secrecy rationale is based on the concern that if patents do not cover all uses, patentees will keep their inventions secret until they find all valuable utilities. There is, however, no need to address this concern with respect to structural biology and genomics because inventions in these fields cannot be kept secret for very long. The human genome has been mapped. Because of the informatics initiatives described above, there is also little danger that structures will not make their way to public databases.[123] As noted previously, it is the norm in the academic community for researchers to deposit their structures in a public database. Although structures found by commercial enterprises are a somewhat different matter, any commercial scientist who publishes must deposit because it is required by journal policies.[124] In addition, the practice of depositing commercial structures appears to one of the authors, who is the Director of the PDB, to be increasing. Besides, an inventor can adopt a secrecy strategy only if she thinks that no one will independently invent the same technology and file for patent protection.[125] Ultimately, high throughput technology makes it so likely that sequences and structures will be found that the secrecy strategy can be safely discounted.

123. Celera's abandonment of its proprietary database model is suggestive. *See* McDonough, *supra* note 68.

124. *See supra* note 48.

125. In jurisdictions where the first to file is awarded the patent, the holder of the secret would become an infringer. Even in the United States, which awards the patent to the first to invent, there is a risk that keeping an invention secret will be regarded as abandonment, suppression, or concealment, and will disqualify the first to invent from obtaining a patent. 35 U.S.C. § 102(g) (2000).

The sufficiency rationale posits that there are cases in which the patented invention has many small markets, each defined by a particular use, and that these must be aggregated to assure that the patentee will receive an adequate return on investment. However, it appears unlikely that biotech inventions fit this description. For some uses, development costs are so low that the recited use will provide a return adequate enough to support the initial investment. Where development costs are very high, the reason will most likely be that the recited use is far upstream. In such cases, recognizing the right provides the patentee with a return that far exceeds its investment.

But while the rationales for broad scope are not applicable, the problems with broad scope are evident. The difficulties in negotiating thickets of rights means that there may be areas that are inefficiently explored. Furthermore, the system may be affording insufficient encouragement to second-comers—to those who find new uses for patented inventions. Of course, such an inventor may be entitled to a patent on the newly discovered use. However, that patent may not be valuable enough to spur adequate research, especially in cases in which research inputs must be licensed. These new uses will likely be protected by process patents, which are difficult to monitor and as a result, generate lower returns. Furthermore, in cases in which the second use infringes the original patent, negotiations for a license may be fraught with difficulties.[126] If following-on is to be encouraged, the system needs to include a way in which pioneer claims can be narrowed.

1. Use Patents

One approach, recently enacted in Germany for genomic patents, is to limit the patentee to the use recited in the patent—that is, to use the *utility* requirement as the measure of *scope*. For example, if the patentee claims that the sequence can be used to diagnose a susceptibility to Condition X, then the patent covers only diagnosis of X. In this way, Germany uses a patent to reward the inventor who found the relationship between the gene and X, but it preserves patent prospects so that others have the motivation to find the gene's other functionalities.[127] Presumably, a similar approach could be adapted for proteins.

126. In theory, the first inventor and second-comer will allocate returns in accordance with their contributions. *See, e.g.*, Suzanne Scotchmer, *Protecting Early Innovators: Should Second-Generation Products Be Patentable?*, 27 RAND J. ECON. 322 (1996). But as we saw earlier, these negotiations often break down.

127. Lord Justice Hoffmann's attempt to narrow the ambit of the EPO patent by confining the claim to "host cells" to cells that do not naturally produce EPO furnishes another example. *See* Kirin-Amgen Inc. v. Hoechst Marion Roussel Ltd., [2005] 1 All E.R. 667 (U.K.).

There are, however, problems with the German approach. This method of narrowing patent scope could set a precedent that will make investors wary of investing capital into new fields. It will require new case law on what counts as an infringing use. It may also create the sort of anticommons problem noted by Rebecca Eisenberg and Michael Heller: Where use of a genomic invention touches on several functionalities simultaneously, it may be difficult to assemble all the rights needed to bring an invention to market.[128] Finally, because this approach would treat genomic and proteomic inventions specially, it may violate the nondiscrimination provision of the TRIPS Agreement.[129]

Explicitly changing the law in this way may also be unnecessary. There is now extensive prior art in the genomics field. If the level of skill were to rise to appropriate levels, it may be that most patentees would not be able to claim products, and would instead be limited by the novelty and nonobviousness requirements to patents on specific processes. In other words, if examination were allowed to proceed on a realistic view of the level of skill in the field, most products would be considered too obvious a patent and claims would only be drawn to particular processes for using the gene or the protein. A result similar to the one in Germany (patents limited to particular uses) would thus be attained under existing law. With this approach, moreover, the problem of defining infringement would be alleviated because the prior art would serve as benchmarks for determining infringement. Furthermore, TRIPS does not set a specific level of skill in the art. Accordingly, this approach also has the benefit of complying with international obligations.

For protein patents that are based on isolation, purification, and crystallization of a protein, scope could also be narrowed by confining the patent to the physical material or, where the material is obvious, to newly found uses. The structural coordinates, which define the physical location of every atom, should be regarded as a description of the protein invention, but not the invention itself. In this way, tool makers (such as designers of microarrays) could use patents to earn a return on their work, but those who merely use data on the shape of folds would not be inhibited. For example, theoreticians would be able to use the information freely for bioinformatics and computational biology purposes.

---

128. Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCIENCE 698 (1998).

129. TRIPS Agreement, *supra* note 86, art. 27(1) (stating that "patents shall be available . . . without discrimination as to . . . the field of technology"); WORLD TRADE ORGANIZATION, CANADA—PATENT PROTECTION OF PHARMACEUTICAL PRODUCTS, WT/DS114/R (Mar. 17, 2000), *available at* http://www.wto.org/english/tratop_e/dispu_e/7428d.pdf. *See generally* Dinwoodie & Dreyfuss, *supra* note 86.

2. Research Exemption

Even if the system is revised as suggested, it is still likely that unique products and processes, critical to life sciences research, will be privatized, and that the rights over these inventions will create obstacles to future creativity. Thus, other ways to limit patent scope may also be necessary. One way to address the problem of chilling research is to attack the problem directly by recognizing a defense in favor of researchers. As noted above, the common law defense has been significantly narrowed; unless the Supreme Court intervenes, the responsibility will fall on Congress. The NAS's study of the patent system discussed various options, including allowing government-funded researchers to use the government's right to practice an invention without authorization.[130] However, the NAS did not, at that time, endorse any particular option.

After the report of that study was published, the American Intellectual Property Lawyers' Association (AIPLA) also suggested the creation of a statutory defense to infringement. Such a defense would immunize the manufacture or use of a patented invention aimed at verifying the validity and scope of the patent; discerning the features, properties, or inherent characteristics or advantages of the invention; finding novel methods of making or using the patented invention; or discovering novel alternatives, improvements, or noninfringing substitutes. In addition, activities incidental to preparations for commercialization of a noninfringing alternative would also be regarded as noninfringing.[131]

This approach, which tracks Judge Newman's dissent in *Integra Lifesciences v. Merck*,[132] and which has quite recently been endorsed by an NAS report on genomics and proteomics,[133] would go a long way to restoring the ability of those in a field to build on the patentee's disclosure. Thus, it would return the system part way to the status quo ante, before the Federal Circuit addressed the experimental use issue.[134] The approach also has the advantage of partially harmonizing U.S. law with the law of other

130. NAS 2004 REPORT, *supra* note 100, at 108–17. A similar idea has been suggested in England. NUFFIELD COUNCIL ON BIOETHICS, *supra* note 84, at 61–62.

131. The American Intellectual Property Lawyers' Association (AIPLA) is apparently now considering a new proposal and has removed the old one from its website. However, its description can be found at AIPLA, *AIPLA Town Meeting on Patent Reform*, 16 IPLS PROC. (Mar. 4, 2005), *available at* http://www.michbar.org/ip/news.cfm.

132. 331 F.3d 860 (Fed. Cir. 2003) (suggesting that a research exemption is important to accomplish the disclosure goals of patent law).

133. NAS 2006 REPORT, *supra* note 14, at 122–23.

134. *See supra* text accompanying notes 25–28.

industrialized countries. Thus, while it is more specific than European law, which typically privileges "experimenting on" the patented invention, the AIPLA proposal would apparently reach the same result in most cases.[135] However, the specificity of the language may give rise to problems of its own because it may make the provision too inflexible to deal with future research strategies.

It is also questionable whether this proposal goes far enough; the experience in Europe is that the comparable rules do not. The exemption permits only uses that focus on learning about the invention that is patented, but researchers often need to use patented inventions for other purposes. Consider, for example, a researcher who relies on a patented diagnostic test to rule out a specific cause of cancer so that she can concentrate on tumors caused by other genetic mutations or a company that tests a new therapeutic agent against a patented therapy. Perhaps these uses would be regarded as incidental activities that are preparatory to commercializing noninfringing alternatives, but it is far from clear that courts would adopt that interpretation. Equally important, the AIPLA formulation does not assure access to research tools or to diagnostic tests when they are used for their intended purpose, and yet experience has shown that these patents are sometimes licensed very restrictively.

An alternative would be to adopt a law similar to that of Japan, which states that "[t]he effects of the patent right shall not extend to the working of the patent right for the purposes of experiment or research."[136] Ostensibly, this is a much broader provision. However, it has not been interpreted by Japan's highest court, and there is reason to think that it will be narrowed to exempt approximately the same activity as is exempt in Europe.[137] After all, an

---

135. Section 11, Number 2 of the German Patent Act provides: "The effects of the patent shall not extend to acts done for experimental purposes relating to the subject matter of the patented invention." Patengesetz [German Patent Act], Dec. 16, 1980 RGBI. 11 at 117, § 11 no. 2 (F.R.G.). Section 60(5)(b) of the U.K. Patent Act provides a defense for acts "done for experimental purposes relating to the subject matter of the invention." Patents Act 1977, c.37, § 60(5)(b) (Eng.) (U.K. law also has an exemption for acts "done privately and for purposes which are not commercial," *id.* § 60(5)(a)). The proposal for a European Community patent also includes an exception for experimental use. For example, article 9(b) of the 2004 draft would have provided that "[t]he rights conferred by the Community patent shall not extend to . . . acts done for experimental purposes relating to the subject-matter of the patented invention." Council of the European Union, *Proposal for a Council Regulation on the Community patent*, File No. 2000/0177 (Mar. 8, 2004).

136. Japanese Patent Act, Law No. 121 of 1959, ch. 4, no. 69(1), *translation available at* http://www.wipo.int/clea/docs_new/en/jp/jp036en.html (last visited Feb. 16, 2006).

137. Thus far, the only cases interpreting the provision is the Japanese case *Ono Pharmaceuticals Co. v. Kyoto Pharmaceutical Industries, Ltd.*, (Sup. Ct., Apr. 16, 1999), *translated summary available at* http://courtdomino2.courts.go.jp./promjudg.nsf/766e4f1d46701bec4925668700435d2e/24e3f24bc3b03e3e49256a93003479d8?OpenDocument (last visited Mar. 9, 2006), holding only that testing during the patent term to obtain data required for regulatory approval was not an infringing act.

overbroad interpretation of a research exemption would be as destructive to the value of patents in research tools in Japan as such an exemption would be here. Since medical research often goes hand-in-hand with diagnosing patients, it could also undermine the value of patents on diagnostics.

The bottom line is that while it is relatively easy to imagine research exemptions that will allow research *on* a patented invention, it is much more difficult to design an exemption that permits research *with* a patented invention but that does not eviscerate incentives to invent research and diagnostic tools. The right result may, therefore, be to take a different approach to research and diagnostic tools and to provide for compulsory licensing in cases in which licensing is inadequate to fulfill social needs.

3. Compulsory Licenses

Because it is unlikely that a research exemption can be fashioned in a way that deals fairly with research tools, diagnostics, and certain other experimental situations, a supplementary approach is needed. Many countries have compulsory licensing provisions that act as safety nets in cases in which there is inadequate access to the patented invention. In Germany, for example, a party who wishes to use a patented invention for the public interest and cannot obtain a license at reasonable cost may apply to the court for a compulsory license. The statute also requires a showing that there is no reasonable alternative to the patented invention, but in cases where the actual human material is necessary, this provision should not present an obstacle.[138] Similar recommendations have been made for the United States. For example, Katherine Strandburg has suggested that a patentee be given a few years in which rights are genuinely exclusive.[139] After that, those who wish to use the patented invention and cannot obtain rights on reasonable terms from the patentee would apply to a court for a compulsory license at a fee determined by the court. Such a procedure would be consistent with the TRIPS Agreement, which envisions the availability of such licenses.[140] Because the patentee could receive a significant return during the exclusive period as well as through the court's order, remuneration would be guaranteed. Indeed, the built-in delay could give the court the information it needs to order adequate royalty payments.

---

138. Bundesgerichtshof [BGH] [Federal Supreme Court], Dec. 5, 1995, 1996 GRUR 190, 1997 IIC 242 (F.R.G.) (revoking a compulsory license because of a lack of sufficient "public interest").

139. Katherine J. Strandburg, *What Does the Public Get? Experimental Use and the Patent Bargain*, 2004 WIS. L. REV. 81.

140. *See* TRIPS Agreement, *supra* note 86, art. 31.

Admittedly, compulsory licensing provisions in patent law have long been resisted in the United States.[141] But it must be kept in mind that genomic and proteomic inventions are significantly different from the inventions of the past. In the usual case, the ability to invent substitutes for the patented invention limits the patentee's ability to hold out. Since the disappointed licensee can find a substitute, there is little need to make compulsory licenses available in cases in which the patentee refuses to deal. However, when the invention is unique—as it is in certain applications of genomic and proteomic patents—another incentive to negotiate is needed. A compulsory license, which would come into play only when the patentee refuses to license on reasonable terms, could serve that function.[142]

CONCLUSION

We are hardly the first to notice that there are problems associated with biotechnology patent law or to suggest that there is something wrong in the way that issues along the subject matter-inventiveness-scope-exemption trajectory have been resolved. In our view, however, reform will not be possible until the relationship between advances in genomics, structural biology, and therapeutics is fully understood. As long as the belief persists that genes lead directly to medicine, there will be a strong incentive to recognize patents far upstream and then to limit them with measures that can sometimes be easily evaded, and are generally counterproductive. We have endeavored to describe the reality of the drug development process, and have also suggested that the move from gene to drug is a long, convoluted process involving risky and creative efforts. These activities yield inventions that are nonobvious even when associated information about the gene or protein is already known. Pushing patenting downstream is thus feasible. It is also desirable as it would increase public access to the fundamental building blocks of scientific knowledge and allow competitive drug development to flourish.

To be sure, there are other ways to cope with upstream patenting. For example, the National Institutes of Health,[143] the Organization for Economic Co-operation and Development,[144] and the National Academies of

141. See, for example, the discussion of compulsory licensing proposals in *Dawson Chemical Co. v. Rohm & Haas Co.*, 448 U.S. 176, 215 & n.21 (1980). Once again, the analogy to copyright is apt as copyright has dealt with the problem of specialized access interests by developing a series of compulsory licenses. *See* 17 U.S.C. §§ 107–122 (2000 & Supp. 2002).

142. Antitrust law may offer another approach to narrowing scope. *See* Dreyfuss, *supra* note 12.

143. Best Practices for the Licensing of Genomic Inventions, 69 Fed. Reg. 67,747 (Nov. 19, 2004).

144. Org. for Econ. Co-operation & Dev., Intellectual Property, http://www.oecd.org/document/50/0,2340,en_2649_201185_34365938_1_1_1_1,00.html (last visited Oct. 31, 2005).

Science[145] have suggested best practices for licensing that would remove some of the friction from the current system. These practices could be made into norms through intervention by these organizations, journals, or even universities themselves. In addition, the National Academies' 2006 Report suggested that one or more of these organizations should develop patent pools or clearing houses to facilitate licensing.[146] In the end, however, we believe that systematic revision is needed. At the very least, the courts must develop a method for dealing with "moving target" issues, such as keeping track of the actual level of skill in the art, the degree to which research functions have been automated, the availability of fundamental data, and the agreements that scientists have made about how such data should be represented.

145. NAS 2006 REPORT, *supra* note 14, at 114–19.
146. *Id.* at 123–24.