Exhibit B

IN THE HIGH COURT OF JUSTICE

CH 1993-K-No. 937
CH 1993 B-No. 4552

CHANCERY DIVISION

PATENTS COURT

In the matter of European Patents (UK) Nos. 148,605 and 411,678 and in the matter of actions for infringement and counterclaims for revocation thereof by inter alia Kirin-Amgen Incorporated, Janssen-Cilag Limited and Roche Diagnostics GmbH

EXPERT REPORT OF PROFESSOR WILLIAM JOHN BRAMMAR

QUALIFICATIONS, EXPERIENCE AND INSTRUCTIONS

1. I have been Professor of Biochemistry in the Department of Biochemistry, University of Leicester, UK, since 1977 and am currently the Dean of the Faculty of Medicine & Biological Sciences. A copy of my *curriculum vitae* and a list of my publications are at Annex 1 to this report.

2. I was trained in Biochemistry (B.Sc.) and Microbial Physiology (Ph.D.) at University College, London, before undertaking post-doctoral work in Biochemical Genetics at Stanford University, California. I was Lecturer in Molecular Biology at the University of Edinburgh from 1967 until moving to the Chair of Biochemistry at Leicester in August 1977.

3. My research at Edinburgh focused on the control of gene-expression in bacteriophage lambda. This allowed me, in collaboration with my local colleagues, Noreen and Kenneth Murray, to make a contribution to the development of derivatives of bacteriophage lambda as vectors for the molecular cloning and *in vitro* manipulation of genes. I was awarded the Colworth Medal for this work by the Biochemical Society in 1977.

4. My research in Leicester has utilised molecular cloning and recombinant DNA techniques to study gene-expression and its control, initially in prokaryotic systems, then moving via lower eukaryotes to higher organisms. I have studied gene-expression in human breast tumours, and systems for manipulating gene-expression in animal and plant cells. My work is now focussed on the molecular biology of ion-channels in mammalian cells and on the control of gene-expression during DNA transfer in bacteria.

5. I have published in the field of gene-cloning since 1975, including the cloning of bacterial genes (1976-1981), fungal genes (1985-1986), mammalian genes (1982-) and cDNAs (1982-). I have written invited reviews on vectors for molecular cloning, and am the co-author of a book on 'Cloning Vectors'. I have also taught undergraduate and postgraduate courses in molecular cloning since 1977.

6. In the late 1970s and early 1980s I described the construction and use of bacteriophage lambda vectors for cloning genomic DNA fragments. Some of these vectors were used in the early 1980s by my laboratory and others to isolate specific genes from mammalian genomic libraries (Mullins *et al*, 1982; Jeffreys *et al*, 1982). In the period 1980-1984, my laboratory was also involved in the construction and screening of tissue-specific cDNA libraries (eg Mullins *et al*, 1982; Windass *et al*, 1984) and in methods for expressing cloned genes in prokaryotic host cells (Moir and Brammar, 1976; Hennam *et al*, 1982; De Maeyer *et al*, 1982).

7. I have previously given evidence at two patent trials in the UK – for Wellcome in the tPA case and for Chiron in the HCV case.

8. I have been asked by the Roche Parties to provide an opinion on a number of matters arising from the United Kingdom legal proceedings between the Roche Parties and the Kirin-Amgen Parties.

9. I have been asked to focus on the molecular biology aspects of the dispute, in particular where these relate to the sufficiency of the Kirin-Amgen Parties' European Patent (UK) No.0148605 ("the '605 patent") and the alleged anticipation by the application for the '605 patent of claim 8 of the Roche Parties' European Patent (UK) No. 0411678 ("the '678 patent").

10. In the course of my report I refer to documents which have been provided to me by the Roche Parties, or that I have located myself. A list of the scientific papers that I refer to in this report can be found in Annex 2 to this report. I also refer to diagrams which were produced to aid the understanding of my report and appear in Annex 3.

CLONING AND EXPRESSING MAMMALIAN DNA IN 1983/84

Introduction

11. In the early 1980s, cDNAs derived from mammalian mRNAs were isolated in order:

(i) to obtain sequences of the encoded proteins, and

(ii) to express in a suitable host cell/vector system in order to produce high yields of the encoded protein.

Genomic clones were usually cloned by groups interested in:

i) the control of expression of the gene, by a study of the promoter region and the proteins interacting with it.

ii) the structure of the gene itself, since the intron-exon pattern could not be predicted from the cDNA sequence.

iii) the relationship between homologous genes within a complex gene-family (eg the members of the β-globin family in higher eukaryotes).

12. The field was initially led by academic groups interested in developing molecular cloning techniques to investigate gene structure and function, but strong teams in established pharmaceutical companies or new biotechnology companies gave great impetus to the field in the early 1980s in their quest to derive processes for the commercial production of valuable proteins.

13. The basic tenets of genetics, the chemistry of DNA, and the processes of gene expression via transcription and translation are dealt with in the draft primer and what follows is merely a brief summary by way of convenience.

14. Genes are transcribed by a family of protein enzymes known as RNA polymerases. The RNA polymerase enzyme unzips the DNA double helix in the region of the gene to be transcribed and then processes along the gene, in a 3' to 5' direction on the template (non-coding) strand, to form a molecule of pre-messenger RNA. The pre-messenger RNA produced from gene transcription must then be modified prior to translation into the gene product by splicing, 5' capping and 3' poly-adenylation to produce mature messenger RNA (mRNA).

15. Not all gene sequences in higher eukaryotes are represented in the mRNA encoding the gene product. The regions that are represented are known as exons, whilst the sequences that intervene are known as introns. In order to form a properly functional mRNA molecule the introns must be excised from the pre-messenger RNA molecule and the exons joined together. This process is known as "splicing" and is performed by complex ribonucleoprotein particles, the spliceosomes, which recognise certain sequences within the pre-messenger RNA molecule as sites to cut the pre-messenger RNA strand and join the strand together to form the mature mRNA (see generally Fig. A, Annex 3).

16. A eukaryotic mRNA will typically contain a translation-initiation codon (AUG), a region which is not translated situated 5' to the initiator codon, an open reading frame which is translated, a stop codon at the end of the coding sequence and a 3' untranslated region with a poly A tail (made up of an adenylate repeat). Sequences in the mRNA itself, particularly in the 3' end of the molecule, can affect the stability of the mRNA molecule in the cell thus increasing or decreasing the efficiency of translation of the mRNA into gene product.

17. The protein or polypeptide produced from the translation of the mRNA can then be further modified inside the cell by the excision of certain portions of the protein, its combination with other polypeptide molecules or the addition of sugars or other molecules in order to produce the fully functional protein end product.

18. The ability to manipulate genes *in vitro* depended on advances in understanding enzymes that work on DNA, including DNA polymerase, DNA ligase and the restriction endonucleases ('restriction enzymes'). Reverse transcriptase, the enzyme that copies an RNA template into a complementary DNA strand, allows the construction and subsequent cloning of DNA coding sequences derived from specific

messenger RNAs, so called cDNAs (copy DNAs). The molecular cloning of cDNAs in plasmid vectors was first described in 1976, and robust techniques for cDNA cloning were widely used by the early 1980s. The 'Maniatis manual' (*Molecular Cloning, A Laboratory Manual*: T. Maniatis, E.F. Fritsch and J. Sambrook: Cold Spring Harbor Press, Cold Spring Harbor Laboratory, New York), a widely used laboratory guide to techniques for cloning DNA molecules, published in 1982, contains a chapter (chapter 7, pp 216-246) on the 'Synthesis and Cloning of cDNA', with detailed practical protocols.

Restriction Enzymes

19. The ability to 'clone' fragments of DNA depended crucially on the isolation of 'restriction enzymes', sequence-specific endo-deoxyribonucleases that can be used to cleave large DNA molecules into discrete, gene-sized pieces. The action of the restriction enzymes can leave the DNA with very short (2-4 nucleotides) regions of single stranded DNA at their ends. As these enzymes are sequence specific, other DNA molecules cut by the same enzyme will have complementary single stranded sequences at their ends, thus aiding the joining or ligation of the two pieces of DNA (see Fig. B(1), Annex 3). These fragments of DNA, produced by the *in vitro* action of purified restriction enzymes, are the objects that are usually cloned. The same restriction enzymes are also utilised to cut the vector DNA to produce ends that are complementary and can be covalently joined to the compatible ends of the DNA fragments using DNA ligase. Restriction enzymes of different sequence-specificities are purified from cells of different bacterial species, where their normal role appears to be the recognition and nucleolytic degradation of 'foreign' DNA.

Vectors

20. Vectors in molecular biology are DNA molecules that are capable of accepting foreign DNA and replicating this DNA in a host organism. The forms of vector DNA we are concerned with in this matter are plasmids and bacteriophage. A plasmid is a circular piece of DNA which is capable of self replication within a host cell and a bacteriophage is a viral particle that infects and replicates within bacterial cells (see Fig. B(2), Annex 3). The way in which foreign DNA is inserted into these vectors is essentially the same. The vector DNA is prepared for the cloning step by cutting it with a restriction enzyme. The foreign DNA of interest can then be cut with the same

restriction enzyme resulting in this fragment having complementary ends to the vector DNA. The foreign DNA and vector DNA are then incubated together with the enzyme DNA ligase and under suitable reaction conditions this enzyme will join the two fragments together to form a recombinant DNA molecule containing both vector DNA and foreign DNA (see Fig B(3), Annex 3).

21. Two classes of vectors used in *E.coli* are plasmids and bacteriophage particles as described above. In the late 1970s and early 1980s, plasmid vectors tended to be used for cloning cDNAs and small DNA fragments, while vectors based on bacteriophage lambda found most common use in constructing genomic DNA libraries. The advantageous feature of vectors based on bacteriophage lambda that makes them attractive for genomic cloning is the process of *in vitro* packaging. Recombinant phage DNAs can be packaged in the presence of concentrated cell extracts containing all the necessary head and tail proteins of the phage, to give fully infectious phage particles. This process is considerably more efficient than transformation, as the process of infection of the *E.coli* host is close to 100% efficient, so that *in vitro* packaging allows recovery of the large numbers of recombinant DNA molecules necessary to constitute a representative library from a complex eukaryotic genome.

Linkers

22. Often short linker oligonucleotides are inserted into vector sequences during cloning procedures. Linkers can fulfil many functions but they are generally carefully designed to have sequences corresponding to desired restriction enzyme cutting sites. Therefore, after the linker has been inserted into the vector sequence the researcher can cut the vector DNA with the restriction enzyme appropriate to the subsequent cloning step. Hence linkers are simply a means of inserting a desired short length of sequence into a DNA molecule.

Promoters

23. Once a DNA molecule has been cloned and characterised, it can be further manipulated for expression to produce the gene-product. Transcription of the coding sequence into RNA requires the host cells' RNA polymerase to recognise an appropriate 'promoter', a sequence motif in the DNA upstream of the coding sequence. *In vivo* interaction of the DNA polymerase with this sequence is modulated

in response to a variety of signals, e.g. metabolic signals, hormonal or other developmental stimuli as a tissue develops. The RNA polymerase/promoter interaction is highly specific and hence it is necessary to provide a promoter that is appropriate for the particular host in the expression strategy. Promoter sequences that work efficiently in *E.coli* were available as portable DNA 'cassettes' by 1983. These could be placed upstream of a cDNA sequence, within an appropriate vector DNA, so that transformed *E.coli* cells would be able to transcribe the cloned cDNA sequence at a high rate. Vectors containing promoters that were known to work in various types of mammalian cells were also available in 1983.

Selective Markers

24. The re-introduction of naked DNA molecules into living cells, the processes known as 'transformation' (for microbial cells) or 'transfection' (for mammalian cells), is very inefficient: the vast majority of the DNA molecules do not enter cells, and most of the cells in the population do not take up the DNA. In order to find the relatively rare cells that have taken up DNA, it is necessary to devise conditions that give such cells a strong selective advantage for growth or survival. For transformation of bacterial cells by plasmid DNAs, the most commonly used selective markers are genes encoding enzymes that inactivate a toxic drug. Transformed cells become drug-resistant, with the rest of the population being killed (or at least failing to grow) in the presence of the drug. Bacteriophage based vectors do not require a selective genetic marker: selection for 'transformation', or infection following *in vitro* packaging, is achieved by the presence of 'plaques', discrete areas of lysis visible as lacunae in a lawn of bacterial cells. Selective markers based on drug-resistance or resistance to an antimetabolite are also available for use with mammalian cells (see e.g. Figs. 20 & 21 of the draft primer).

Constructing cDNA libraries in 1983/84

25. The cloning of cDNAs derived from mammalian mRNAs was first achieved in 1976. In the period up to 1983 significant improvements in the techniques of cDNA synthesis and cloning had been achieved, making the processes more efficient and more likely to give fully representative cDNA libraries. cDNA cloning procedures utilise the entire population of mRNAs *expressed* in a given tissue or cell-type. To favour isolation of desired cDNAs, cells that express the appropriate gene at high

rates are chosen as the source of template mRNA. The targeted cDNAs will then tend to represent the same proportion of the total clones in the library as the specific mRNA species represents of the mRNA population. In very favourable cases, this can be 1% or more.

26. The construction of a cDNA library is achieved by first identifying a tissue or immortalised cell line which expresses the protein of interest. This ensures that the mRNA for the gene of interest will be expressed and will form part of the total mRNA isolated from the tissue or cell sample. Total mRNA is then isolated from the sample and a DNA copy is made of these molecules by the enzyme reverse transcriptase (see Fig. C generally). A poly dT primer is used for this reaction as this will anneal to the polyA tail of the mRNA molecules. Reverse transcriptase produces a single stranded complementary DNA copy of the mRNA, that can be extended using deoxynucleotide terminal transferase at its 3' end. The synthesis of the second strand, using DNA polymerase, can be primed with oligo dG (or oligo dC) and produces a double stranded cDNA. The double stranded DNA molecules that are produced in this way are tailed with homopolymeric bases, e.g. polydG or polydC, and then ligated to prepared vector molecules with the complementary homopolymeric tail to assist the efficiency of the cloning procedure. The recombinant plasmid DNA is then used to transform bacteria in order to be amplified and, subsequently, screened in order to identify the recombinant DNA of interest.

27. A development in which the synthesis of first-strand cDNA is primed from an oligodT tail covalently attached to one end of the linearized vector DNA, devised specifically to facilitate the cloning of full-length cDNA (Okayama and Berg, 1982), became widely accepted and used by 1983 and was the method used by Amgen to construct cDNA libraries from the kidney mRNA of anaemic monkeys.

Constructing Genomic Libraries in 1983/84

28. The construction of DNA libraries representing the genomic DNA from higher eukaryotes, including mammals, became possible through the development of vectors based on bacteriophage lambda in the late 1970s. As I have already mentioned, the particular advantage of lambda as a vector was the ability to package recombinant DNA into a bacteriophage particle *in vitro*, a process that was considerably more efficient than transformation with unpackaged DNA. This relatively efficient

recovery of recombinant molecules allowed the construction of the many (*ca* 1 million) recombinants that are required to give a good representation of the entire mammalian genome in a library. Papers giving detailed procedures for constructing representative libraries of mammalian genomes were published by the Maniatis laboratory (Maniatis *et al*, 1978; Lawn *et al*,1978). Amgen's isolation of human genomic clones was carried out with a donated sample of the 'Lawn' human genomic library.

29. Briefly, this technique involves extracting DNA from a tissue source and cutting it with a restriction enzyme after which DNA of the required size (about 15-20 kilobases) is then isolated prior to cloning (see generally Fig. D, Annex 3). The vector is prepared by cutting the bacteriophage lambda chromosome with a restriction enzyme which recognises a compatible sequence to that used to cut the genomic DNA. The "arms" of the lambda chromosome are then isolated and the middle portion of the chromosome discarded. These arms are then incubated with the fractionated genomic DNA in a ligation reaction in order to produce recombinant DNA molecules. The resulting recombinant molecules can then either be transformed into bacteria (a low efficiency procedure) or packaged into infective bacteriophage particles by mixing them with the bacteriophage head and tail proteins prior to infecting *E.coli* (a high efficiency process). Infective bacteriophages will spontaneously form and these bacteriophage particles can then be used to infect *E.coli*. These plaques will then have to be screened to identify the recombinant clone of interest.

Isolating the cDNA or genomic DNA From the Library

30. A crucial stage in the molecular cloning of a particular DNA sequence is the detection of the specific recombinant DNAs within the library. The techniques used to screen cDNA and genomic libraries are similar and the most commonly used method involves the process of nucleic acid hybridisation, in which nucleic acids of complementary sequence will find each other and anneal by hydrogen bonding. If one of the two reacting species (the 'probe') is tagged for easy detection, usually by labelling with a radioactive isotope (e.g. ^{32}P), it will allow sensitive detection of molecules in the clone library that contain the complementary sequence.

31. The use of labelled nucleic acid probes of this kind depends on information about the target sequence. In the absence of prior cloning, the necessary sequence information can be derived from the amino acid sequence of the protein-product of the relevant desired gene. The protein is purified from the most appropriate source and the sequence of amino acids at the N-terminus of the pure protein can then be determined by successive rounds of chemical degradation. Also, the protein can be cut up using enzymes such as trypsin into small fragments, which are then separated and sequenced from their N-termini.

32. The general principles of amino acid sequence determination have not changed over many years. Instrumentation has improved, however, leading to increases in speed and sensitivity. In particular, the introduction of the gas-phase 'sequenator' in 1981/1982 brought a significant increase in sensitivity, allowing sequence to be determined on picomole quantities of protein or peptide. (The increase in speed and sensitivity of the gas-phase sequenator over its predecessors is based on a design that allows the cleaved N-terminal amino acid derivative to be delivered to the detection system in the gaseous phase, rather than in liquid solution).

33. Once a short segment of amino acid sequence has been determined, the mRNA sequence encoding the amino acids can be predicted from a knowledge of the Universal Genetic Code thus enabling the design of suitable oligonucleotide probes. The code is 'degenerate', with more than one codon specifying most amino acids, and as many as six encoding leucine, serine and arginine. Amino acid sequences containing amino acids that have only one codon (ie tryptophan and methionine) or two codons (there are nine amino acids with this pattern) are favoured as sequences on which to base the design of probes: amino acids with four or six codons are avoided, if possible. (Peptides containing tryptophan residues can be detected by their relatively high absorbance at 280nm, allowing the sequencing effort to be focussed on these peptides.) Oligonucleotides containing at least thirteen nucleotides are usually used as probes, in order to generate the required specificity, though oligonucleotides as short as eleven nucleotides have been successfully used for screening cDNA libraries. (Edlund *et al*, 1983: Suggs *et al*, 1981). Because of the unavoidable ambiguity due to codon degeneracy, all possible variants of the coding sequence can be synthesised as a 'mixed pool'. One of the variants in the pool will be correct, and should be capable of acting as a specific probe for the desired coding sequence.

34. The 'Maniatis manual' also discusses methods for the identification of specific recombinant clones, by methods based on molecular hybridisation of nucleic acids or on genetic recombination. The use of mixed oligonucleotide probes, designed on the basis of amino acid sequences derived from the protein-product of the relevant gene and the Universal Genetic Code, is described as "the method of choice" for screening cDNA libraries if partial or complete amino acid sequence of the protein of interest is available (p.226). Practical details on the use of mixed oligonucleotide probes are not provided by the Maniatis manual, because the author's laboratories had not used these methods at that time, but the procedures were well established and widely used by the end of 1983.

35. Once the probes have been designed and made, the next step is to screen the library (see generally Fig. E, Annex 3). Transformed bacteria, grown as isolated colonies on selective plates, or bacteriophage plaques growing on a bacterial lawn, are transferred to a solid matrix, usually a nitrocellulose filter. (The original 'master' plates are retained for the subsequent isolation of viable bacteria or phages). The bacterial cells are lysed *in situ* and the DNAs fixed to the filter by 'baking' in a vacuum oven. The filters are then immersed in liquid and the probe mixture is applied under conditions designed to allow hybridisation. The temperature, ionic strength and time to be used can be roughly calculated, knowing the length and G/C content of the least strongly hybridising component of the probe mixture. Following washing of the filters, to remove non-specifically attached probe, and gentle drying, the filters are exposed to X-ray film to reveal DNAs that have hybridised to the probe. The length and stringency of the washing can be adjusted empirically to eliminate non-specific hybridisation to the 'background' majority of the DNAs. Because a pool of oligonucleotides is used as a probe, the conditions of hybridisation and washing cannot be too stringent and positively hybridising clones are simply candidates which may contain the sequence of interest.

36. Mixed oligonucleotide probes have to be used with caution, under conditions of temperature and ionic strength that do not preclude hybridisation of that component of the mixture calculated from its base composition to give the weakest hybridisation. (The higher the G/C content, the higher the melting temperature of the DNA hybrid.) Under such conditions, it is inevitable that the mixed probes will give rise to significant numbers of false positives. Where large numbers of positive clones are

detected, the number of false positives can be reduced by screening with a second, independent probe based on a different region of amino acid sequence. Isolates that are positive with two independent probes can then be checked for authenticity by direct nucleic acid sequence determination. Clones encoding human tissue plasminogen activator (Edlund *et al,* 1983) and a sub-unit of the nicotinic acetyl choline receptor from the Californian ray (Noda *et al*, 1982), for example, were isolated using this strategy.

cDNA Versus Genomic Route to Cloning

37. Although the procedures for screening cDNA and genomic DNA libraries by hybridisation to a labelled nucleic acid probe are similar, the two types of library differ in both their content and the nature of the information that can be obtained from them. The cDNA library has been prepared from messenger RNA as the starting material: only expressed sequences are represented in the library, and the frequency at which any coding sequence is represented is roughly proportional to the relative rate of expression of the cognate gene. In practice, this varies over a wide range, from *ca.*0.001% to *ca.*1%. For a genomic DNA library, all single-copy genes should, ideally, be equally represented, and the DNA included in any one recombinant bacteriophage represents about 0.0005% of the sequence information in a mammalian genome. This means that one usually has to screen larger numbers of recombinants to find positive clones in a genomic DNA library than in a cDNA library. How many more depends on the rate of expression of the relevant gene in the tissue from which the cDNA library was constructed.

The advantages of using a cDNA clone to achieve the expression of a polypeptide

38. A cloned cDNA suitable for expression in a heterologous cell will contain a 5' non-coding region, a translation-start codon, an open reading frame encoding the protein of interest, a translation-termination codon and a 3' untranslated region. It would readily be rendered expressible by placing it downstream of an appropriate promoter and transcription-start site in a vector. In some cases, it is also advantageous to arrange for the vector to provide an intron and a polyA-addition sequence downstream of the coding sequence.

39. In contrast, the genomic DNA is cloned as a 'random' fragment of DNA that contains the introns and exons of the gene, the promoter, enhancers and other regulatory elements, as well as genomic DNA that may well not relate to the gene of interest. To arrange for expression from a promoter that is known to work well in the intended host cell, the genomic DNA must be trimmed in such a way as to eliminate the native promoter and place the first exon closely downstream of the transcription-start site. Effective expression requires that the transcript is properly processed in the heterologous host cell and that the mRNA produced can be efficiently translated. The splicing step can be problematical, because the patterns of splicing of some genes can vary between cell-types.

40. The relatively small size of the cDNA also provides advantages over the usually much larger genomic DNAs. (EPO is not unusual in having only about 10% of the gene as coding sequence). The fact that the primary transcript is much shorter is an energetic advantage to the cell that is expressing a heterologous gene at high rates, leading to increased stability and faster growth rates. Because the cDNA is much shorter than the genomic DNA sequence, it is likely to be much easier to manipulate for the purposes of expression or designed alteration. *In vitro* mutagenesis, for example to improve the efficiency of translation by altering the sequence ('Kozak consensus sequence') adjacent to the translation-start codon, or to change the coding sequence in order to alter the properties of the encoded protein, is much easier in the context of a cDNA sequence than in that of the relatively complex genomic DNA sequence. In addition, plasmid vectors with high copy numbers, such as those used in the '605 patent, are much more unstable in *E. coli*, the standard host for their propagation, when they contain the relatively large DNA inserts characteristic of genomic DNA coding sequences.

41. A cDNA clone is also desirable as the best characterised gene expression systems in use in 1984 were *E.coli* and yeast systems which do not permit expression of complex mammalian genomic DNA sequences due to the presence of introns.

42. Genomic clones will contain introns as well as exons. We know from post-1983 work that some mammalian genes are too large to clone in single plasmid, phage or cosmid vectors. In such cases, the full length gene may have to be reconstructed from a series

of overlapping cloned segments prior to expression in a suitable eukaryotic host cell system.

43. In 1983/84 it was apparent that cDNA cloning was the preferred route for obtaining coding sequences to express for the purpose of protein production. The relative simplicity of the cDNA, compared with the gene itself, the comparative ease of manipulating the sequence for expression, and (usually) an enriched source of the relevant mRNA all favour the cDNA route. The difficulty facing the experimentalist in 1983/84 who wanted to clone a human EPO coding sequence was that no known source of human EPO mRNA was available.

The skilled person

44. In 1983/84, isolation and manipulation of specific cDNAs and genomic clones would require a molecular biologist with post-doctoral experience. Protein purification and sequencing would also require a similar level of expertise in the separate discipline of 'protein chemistry'. Bringing clones to successful expression in mammalian cells would require expertise in cell-culture, which might be provided by a skilled technician under the general direction, say, of the molecular biologist. Such a team would also probably have the services of a technician who would be involved in sample preparation and 'routine' analysis, and who might, for example, have carried out or helped with DNA sequencing. This suggested team of four, at least two at post-doctoral (Ph.D) level, would be the minimum required to progress the project at an acceptable rate.

THE '605 PATENT

The background

45. The background to the invention of the '605 patent is explained in two sections spanning pages 2 to 7. Under the heading "Manipulation of Genetic Materials", the inventor introduces recombinant DNA technology and at p.4 lines 6-16 sets out three alternative methods for obtaining DNA to insert into a vector. The three methods identified are:

(1) isolation of DNA from genomic DNA;

(2) synthesis of a DNA sequence coding for the polypeptide of interest; and

(3) reverse transcription of mRNA isolated from cells, to produce cDNA.

The inventor is right – these were the three distinct ways in which DNA could potentially be obtained for insertion into a vector.

46. At p.3 lines 17-28 the inventor suggests that synthesis is frequently the method of choice when the whole amino acid sequence of the polypeptide is known.

47. Although my area of specialisation is molecular biology, I followed developments in the field of oligonucleotide synthesis with great interest. My laboratory was sponsored by the international chemical and pharmaceutical company, ICI, from 1979 until 1992. In the early years of that arrangement, I had close contact with the company's synthetic organic chemists, Drs Mike Edge and Alex Markham, who led the team that synthesized a synthetic sequence encoding human (leukocyte) α-interferon (Edge *et al*, 1981). This project involved an intensive effort over many months by an expert team of six chemists. Some steps in the synthesis of the intermediate oligonucleotides did not proceed satisfactorily and had to be redesigned. The final synthetic coding sequence was cloned into a specially constructed *E.coli* expression vector in my Leicester laboratory and used to produce 'recombinant' interferon in *E.coli* (De Maeyer *et al*, 1982) and in the methanol-utilising bacterium, *Methylophilus methylotrophus* (Hennam *et al*, 1982).

48. In my experience, in 1984, chemical synthesis of DNA of sufficient length to code for polypeptides of the length of EPO was very difficult and inefficient and it would not have been the method of choice.

49. The inventor then deals with obtaining cDNA coding for the polypeptide of interest (p.3 line 29 – p.4 line 5). The inventor suggests that this was the method of choice if synthesis were not possible. As I have previously mentioned, in my view, it was the method of choice provided a source of mRNA was available. The inventor then discusses the methods available for probing cDNA libraries and, in particular, the use of mixed oligonucleotide probes. As the inventor says (p.3 lines 49-55) the use of mixed oligonucleotide probes in cDNA libraries was particularly useful in detecting cDNA clones derived from sources which provide extremely low amounts of mRNA for the polypeptide of interest.

50. The inventor then cites various publications relating to the use of mixed oligonucleotide probes in cDNA libraries and says that the technique (first described by Wallace *et al*, 1981) had advanced to the point where success had been achieved using a pool of 32 16-mers and a single 11-mer probe from a different region. Other examples could have been given: Woods *et al* (1982) had isolated cDNA clones coding for human complement protein factor B using two fully degenerate pools of 17-mers corresponding to different regions of the sequence, one pool of 32 probes and one pool of 48 probes; Whitehead *et al* (1983) had isolated a cDNA clone coding for human complement C4 using a fully degenerate pool of 384 23-mers. For that reason, the inventor is wrong to claim, as he does at p.11 line 3, that the '605 patent is the first known instance of a description of the use of a mixture of more than 32 oligonucleotide probes in the isolation of cDNA clones. The use of mixed oligonucleotide probes to screen cDNA libraries was well established by the end of 1983, as the inventor recognises at p. 3 line 49 – p. 4 line 5.

51. The inventor then turns to genomic routes (p.4 lines 6-33), saying that this approach was the least common of the three.

52. It seems from the paragraph starting at line 34 of page 4 that the inventor is looking for improved methods for isolating mammalian DNA in circumstances where *"...little is known of the amino acid sequence of the polypeptide coded for and where "enriched" tissue sources of mRNA are not ready available for use in constructing cDNA libraries"*.

53. The next section is headed "Erythropoietin as a Polypeptide of Interest". After discussing the problems of isolating EPO in sufficient quantities for use in therapy from natural sources, the inventor turns to the attempts that have been made to isolate DNA coding for EPO (p.6 line 57 – p.7 line 28). The inventor says that none have been successful:

> *"...due principally to the scarcity of tissue sources, especially human tissue sources, enriched in mRNA such as would allow for construction of a cDNA library from which a DNA sequence coding for erythropoeitin might be isolated by conventional techniques. Further, so little is known of the continuous sequence of amino acid residues of erythropoeitin that it is not possible to construct e.g. long polynucleotide probes readily capable of reliable use in DNA/DNA hybridisation screening of cDNA and especially genomic DNA libraries."*

54. In other words, the inventor is identifying two problems standing in the way of obtaining a DNA coding for human EPO. The first, and most important, was the absence of a suitable tissue source; without such a tissue source a suitable cDNA library could not be constructed. The second was the limited amount of sequence information that could be used to design probes to probe either a genomic DNA or cDNA library.

55. The inventor then deals with the most successful attempts to date to provide DNA coding for EPO. The work of Farber *et al,* (1983), (p.7 lines 11-23) involved *in vitro* translation of mRNA derived in one case from the kidneys of phenylhydrazine treated baboons and in the other case from human kidney. Farber had suggested that this might allow for the construction of a human kidney cDNA library from which EPO cDNA might be isolated. However, this suggestion does not seem to have been taken up by the inventor.

The inventor's work

56. The inventor describes what he did to overcome this problem in the detailed description starting at p.11. The first step (see Example 1A on p.12) was to isolate human EPO from urine and subject it to tryptic digestion. 17 tryptic fragments were obtained in amounts of about 100-150 picomoles. In fact, I understand that these fragments were obtained in an already separated form from a consultant to Amgen, Dr Eugene Goldwasser. The tryptic fragments were then sequenced in a gas-phase sequencer. This was a standard procedure by 1983.

57. As far as I can see, the inventor has dealt with the problem of obtaining sequence information about EPO by simply obtaining sufficient quantities of EPO tryptic fragments and sequencing them.

58. The sequence information was then used (Example 1B) to design probes in the manner in which I have described in paragraph 33 above. I note that some of the sequence data obtained from the tryptic fragments were particularly useful for designing sets of probes as there is a high occurrence of amino acid residues coded for by either one or two codons. For example, in peptide T35 the sequence valine (V, 4 codons), asparagine (N, 2 codons), phenylalanine (F, 2 codons), tyrosine (Y, 2 codons), alanine (A, 4 codons), tryptophan (W, 1 codon), and the first two nucleotides

of the lysine (K) codon (which are identical in both the codons for lysine) can be used.

59. The inventor used sequences derived from peptides T35 and T38 to design probes as these fragments both have peptide sequences that flank tryptophan (W) residues: these peptides would have been readily identifiable by their relatively high absorbance at 280 nm. The probes (a pool of 128 20-mers and a pool of 128 17-mers from different regions of the sequence) were then synthesised.

60. If a source of human EPO mRNA had been available, the inventor could have used these probes to screen a cDNA library made from that mRNA. Instead, the inventor used his probes to probe a monkey cDNA library and a human genomic library.

61. The monkey cDNA work is reported in Examples 2 and 3. Monkeys were made anaemic using phenylhydrazine. This treatment was necessary because the EPO gene is inactive under normal conditions, when the hormone is not required, but is activated into expression by oxygen deficiency. Phenylhydrazine causes the latter condition by inhibition of the production of red blood cells, leading to their gradual depletion in the circulation of the treated animal.

62. The monkeys' sera and kidneys were then harvested (Example 2A). It goes without saying that this procedure cannot be carried out on humans! The sera of the treated monkeys was shown to contain elevated levels of EPO (Example 2B). The inventor then isolated mRNA from the monkey kidneys and made a cDNA library using the Okayama & Berg procedure (Example 3A). He then screened the library using probes he had designed (Example 3B). Seven positive clones were found, one of which was sequenced (Example 3C). The sequence obtained is set out in Table V together with the deduced amino acid sequence.

63. The human genomic library work is reported in Examples 4 and 5. The inventor used a genomic library made by Lawn. He screened the library using the probes he had designed, revealing three positive clones (Example 4B). One of the clones was sequenced (Example 5) and the sequence obtained is set out in Table VI. Table VI also shows the deduced amino acid sequence together with the deduced locations of the introns that interrupt the coding sequence (deduced by comparison with the monkey cDNA and from knowledge of putative splice sites). Example 5 finishes by

referring to a comparison of the deduced amino acid sequences of monkey and human EPO and to the presence of an additional lysine residue in the human sequence.

64. Examples 6 and 7 relate to the creation of expression vectors. Example 6 concerns the construction of a monkey cDNA expression vector while Example 7 is about constructing human genomic DNA expression vectors. The expression vectors are then used to transfect COS cells (Examples 6 and 7) and CHO cells (Example 10). I shall deal with Examples 6, 7 and 10 in more detail below. Examples 11 and 12 relate to the construction and expression of synthetic DNA sequences designed for introduction into *E. coli* and yeast and are dealt with in the expert report of Professor Gassen.

65. So, to summarise, what the inventor has done is to avoid the problem of the absence of a suitable human tissue source for construction of a cDNA library by instead isolating DNA from human genomic sources. In the case of the monkey, he has generated a suitable tissue source by making the monkey anaemic. He has not solved the problem of how to obtain human EPO cDNA.

HUMAN EPO cDNA

66. I have been asked to comment on whether I think the '605 patent would enable a team of ordinary skill in the art to isolate a human cDNA sequence encoding biologically active human EPO. As I explained earlier, the '605 patent solves one of the perceived problems preventing the cloning of EPO cDNA, i.e. the lack of sequence information. In my opinion the '605 patent solves this problem in two ways: first by giving the sequence of the tryptic fragments T4a→T38 (Table I, Example 1), and second, by disclosing the genomic EPO sequence (Table VI, Example 5). Both of these sources of sequence information could be used to design probes to isolate a human cDNA encoding EPO from a cDNA library. However, in order to make a suitable cDNA library a source of human EPO mRNA is still needed.

67. I will consider in turn whether:

(a) the '605 patent discloses a tissue source producing human EPO mRNA; and

(b) the sequence disclosed by the '605 patent can be used to create a source of human EPO mRNA by transfection into COS cells.

68. I understand that Professor Gassen will consider in his expert report whether a sequence identical to a cDNA sequence could be synthesised following the teaching of the '605 patent.

Tissue Sources of Human EPO mRNA

69. As is acknowledged in the '605 patent, the major problem in obtaining a cDNA encoding human EPO was the lack of a known and obtainable tissue source enriched in human EPO mRNA such as would allow for construction of a cDNA library from which the human EPO cDNA sequence might be isolated.

70. As I mentioned above, the '605 patent cites at p.7 the work done by Farber et al. confirming the kidney as the site of human EPO expression. Although EPO is produced in the adult kidney, the gene is only significantly expressed under conditions of oxygen limitation. For obvious reasons one cannot induce anaemia in human subjects in order to trigger the production of EPO as Amgen did in monkeys. Thus obtaining human kidney expressing the EPO gene was impractical and the '605 patent does not suggest any other tissue source from which a cDNA library could be made.

The COS Cell Route

71. My attention has been drawn to the reference in Example 5 of the '605 Patent to human EPO cDNA that reads (p.25, lines 17-20):

> *"Presence of the lysine residue in the human polypeptide sequence was further verified by the sequencing of a cDNA human sequence clone prepared from mRNA isolated from COSP-1 cells transformed with the human genomic DNA in Example 7, infra".*

72. This passage appears in the context of a comparison made between the monkey and human DNA sequence. There is no suggestion that this is a method of making an authentic human cDNA and no characterisation of the isolate, except that it encodes a lysine at position 116.

73. The positions of the exons and introns in the human EPO gene were determined in the '605 patent by comparison between the monkey cDNA sequence and the human genomic EPO sequence. The sequence similarity between monkey and human EPO is

sufficiently high to allow this comparison, but the conclusion would have been firmer and more direct if the inventor had been able to compare a human cDNA sequence with the human genomic sequence. Furthermore, had the inventor obtained an authentic human cDNA clone, the skilled reader would have expected its sequence to be disclosed in the '605 patent.

74. Accordingly, I do not believe that the skilled reader of the '605 patent would regard this passage as a suggestion to use transfected COS cells to produce a human cDNA clone.

75. Even if the '605 patent were suggesting a route for making a human cDNA sequence via expression of the human EPO gene in COS cells, such a route would be thought to be fraught with risk of artefact and would have been considered bad practice, if it had been considered at all. As I mention below, later experimentation has proved these fears to be justified.

76. The COS cell line was originally specifically constructed to allow independent replication of plasmid DNA sequences containing a region of DNA sequence known as the SV40 origin of replication *(ori)*. This DNA sequence is derived from the genome of the SV40 virus. *In vivo* the *ori* sequence is essential for replication of the SV40 genome in eukaryotic cells and is also utilised to drive DNA replication when included in plasmid DNA.

77. The progenitor of COS cells is the CV-1 monkey kidney cell line. This cell line is a 'permissive' host for infection and killing by SV40, that is it is susceptible to infection by, and the subsequent growth and replication of, the SV40 virus. The SV40 T-gene encodes a protein, the 'T-antigen' or 'large T-antigen', that is required to activate SV40 DNA replication from the *ori* sequence. The 'T-antigen' activates replication by binding as an oligomer at the *ori* sequence and initiating the replication process. The COS cell-line was constructed by integration of cloned SV40 genes, including the T-gene but lacking *ori*, into the genome of CV-1 cells. The resulting cell-line, COS (cv-1, origin, SV40), produces the T-antigen and all other factors necessary to activate and maintain DNA replication initiated from the SV40 origin.

78. Plasmid DNAs containing the *ori* sequence are immediately driven to replicate, on entering the COS cells after transfection, by the resident SV40 T-antigen. This rapid

and continuous replication, termed 'runaway replication', leads to hundreds of thousands of copies of the plasmid DNA per cell within two days. This very high gene dosage leads to a short-term burst of synthesis of the encoded proteins before the host cell dies from energy depletion and other deleterious effects.

79. The COS cell system has proved useful as a laboratory tool for analysing the production and post-translational processing of mammalian proteins. But mammalian cells are not geared to the expression of 200,000 gene copies from a strong promoter. The cell's capacity for pre-mRNA processing, particularly the splicing reaction, is overwhelmed, leading to artefacts at a very high frequency (Wise *et al*, 1989; Nordstrom & Westhafer, 1986; Laub *et al.*, 1982).

80. An illustration of the problems that are likely to arise in attempting to use COS cells transfected with plasmids containing EPO genomic clones as a source of EPO mRNA from which to make cDNA is found in the 1987 publication of Wojchowski *et al.* These authors attempted to use the transfected COS-1 cell route to make EPO cDNA for expression in other cell types, particularly insect cells. They started by cloning the EPO gene in a λ vector, probing the genomic library with a unique oligonucleotide designed from the sequence data from the Jacobs *et al*, (1985) publication. The EPO gene was excised on a *Bst*EII-*Bam*HI fragment, as in Example 7B of the '605 patent, and placed in an expression vector (pSV2) for COS cells, downstream of an SV40 promoter. The essential features of the expression construct are similar to those used in Example 7B of the '605 patent (see Fig. F(1), Annex 3).

81. This construct was introduced into COS-1 cells, polyA$^+$ RNA was isolated and used to construct a cDNA library in the λ vector, λgt10, using *Eco*RI linkers. Phages containing EPO cDNA sequences were recognised with an EPO genomic probe and EPO cDNA inserts were characterized by DNA sequencing. The isolated EPO cDNA clones were shown to be incomplete at the 3' end, where EPO coding sequence from exonV was fused to SV40 DNA sequences derived from the expression vector. As far as one can tell from the publication, part of the downstream EPO coding sequence at the end of exon V seems to have been spliced out.

82. Wojchowski *et al.* then devised a strategy for completing the EPO cDNA coding sequence, using exonV sequences previously cloned as genomic DNA. The steps

involved in this 'correction' procedure are indicated in Fig. F(2), Annex 3. The procedure is complex and involves the following steps:

(i) A small *Sac*I fragment of the EPO gene in its plasmid vector ('gEPO^{1}') was removed by cutting with *Sac*I and religation, in order to eliminate an unwanted *Acc*I target-site in the construct (Fig. F(2), step1).

(ii) The *Bssh*II to *Acc*I fragment (2.12 kb) of the resulting plasmid, 'gEPO^{2}', containing the EPO coding sequence except for the final part of exon V, was removed (Fig F(2), step 2).

(iii) A *Bssh*II to *Acc*I fragment (540 bp) of the EPO cDNA plasmid was isolated (step 3) and ligated into the 'arms' of gEPO^{2}, creating the reconstituted genomic/cDNA chimaera without introns and with a 'repaired' coding sequence 3' to the *Acc*I site in the COS-derived EPO cDNA.

83. The resulting construct is thus a hybrid between genomic clone and cDNA with the structure indicated in Fig. F(3), Annex 3.

84. An intronless coding sequence was then isolated by excision with *Bss*HII and *Bgl*II and introduced into appropriate expression vectors using *Bam*HI linkers.

85. Similar artefacts to those described by Wojchowski et al. have been seen with other sequences cloned into SV40-based expression vectors, frequently in COS cells where the copy number of the vector is so unnatural. These aberrations are very common and are usually due to aberrant splicing events between cryptic splice-sites in coding sequences and a downstream intron in the vector (e.g., Wise *et al.*, 1989; Nordstrom and Westhafer, 1986).

86. In conclusion, I do not think there is anything in the '605 patent that would suggest to the skilled reader to use transfected COS cells to produce a human EPO cDNA clone. If such a route were indeed considered and embarked upon, the use of the transiently transfected COS cell system for generating authentic, mature mRNA via forced expression of a mammalian gene is far from straightforward. Although some correctly processed mRNA molecules are likely to be produced, these will be part of a population of aberrantly processed transcripts. Distinguishing the authentic from the aberrant will often be a significant problem and artefacts will occur with high frequency.

PUTTING THE '605 PATENT INTO PRACTICE

87. I have been asked to consider whether a skilled reader of the '605 patent could have followed its teachings so as to obtain a recombinant human EPO.

88. I also understand that the Kirin-Amgen parties are arguing that the application for the '605 patent anticipates claim 8 of the Roche Parties' '678 patent. Accordingly, I have been asked to consider whether the description in the '605 patent application is either inaccurate or incomplete with regards to the production of CHO cells which contain, operatively linked to an expression control sequence, a DNA sequence encoding human EPO. I am told that Examples 6 and 7 and (in all respects which are relevant to what I have to consider) Example 10, are identical in the '605 patent and in the application, so I just refer to the patent below.

89. For the reasons listed below, the descriptions in the examples of the '605 patent would have presented the skilled reader with a number of problems that would have required significant effort and burden to overcome.

Example 4

90. This example relates to the creation of a human genomic library and the procedures to screen it for a DNA sequence encoding EPO.

91. The human genomic DNA library used by the inventor was prepared by the Maniatis laboratory at Harvard (Lawn *et al*, 1978), using a bacteriophage vector, λCh4A, prepared by Blattner's laboratory at Wisconsin (Blattner *et al*, 1977).

92. Although procedures for creating genomic libraries in lambda phage vectors from mammalian DNAs were fully described by Maniatis *et al.* (1978) and Lawn *et al.* (1978), they were technically demanding and time-consuming to repeat. Human DNA, after controlled partial digestion with *Hae*III and *Alu*I, both of which cut frequently and produce blunt termini, was selected from the 15-20 kb range by sedimentation in a sucrose gradient. Internal *Eco*RI sites in the genomic DNA were protected by modification with *Eco*RI methylase before terminal addition of synthetic linkers containing *Eco*RI target sites. After *Eco*RI cleavage, the DNA fragments were again size-fractionated by sedimentation in the ultracentrifuge to eliminate excess linkers that would compete unfavourably in the ligation to vector DNA. The

recovered DNA was ligated into the purified, *Eco*RI-generated 'arms' of the vector phage. *In vitro* packaging into phage particles finally led to the creation of the desired library. A number of the necessary reagents for the procedure, including *Eco*RI methylase and packaging extracts were not commercially available and had to be prepared in the laboratory. Simply making satisfactory and useable human genomic libraries by these procedures would have been a daunting task for most laboratories that would have occupied several man-months in 1983.

93. Many other laboratories in the US and Europe, including my own, had constructed their own genomic libraries in bacteriophage or cosmid vectors and isolated specific genes of interest by adapting the described procedures. (Mullins *et al*, 1982; Jeffreys *et al*, 1982: Ish-Horowicz & Burke; 1981). These libraries were made by different procedures and from different sources of genomic DNA potentially containing different allelic variants of the genes and would not have been the same as the Lawn library.

94. The Lawn library used by Amgen to isolate human EPO was generally regarded as being of excellent quality and was therefore in great demand. By 1983, however, the Lawn library was exhausted and aliquots from it were no longer being provided.

Example 6

95. This example relates to the creation of plasmid pDSVL1 for use in the expression of monkey cDNA encoding EPO in COS cells. This example is, however, of relevance to the sections dealing with the expression of human EPO sequence as pDSVL1 is also used to express human genomic clones encoding EPO in COS and CHO host cells in Examples 7B and 10 respectively.

96. pDSVL1 was constructed "in-house" to allow for propagation in *E.coli* and mammalian cells. Figure G, Annex 3 is an attempt to understand and explain the construction of pDSVL1, though there remains a considerable element of doubt associated with some of the steps.

97. pDSVL1 comprises the following components:

(i) A pBR322 fragment comprising nucleotides 2448 through 4362. This fragment contains the origin of replication for *E.coli* and the selectable ampicillin-resistance

gene. Although the nucleotides that comprise the pBR322 fragment are specified, there is no teaching in the '605 patent nor is it obvious to me how this fragment was obtained. In particular, there is no indication of how the cut at nucleotide 2448 necessary to obtain this fragment was achieved as no restriction site exists which allows the DNA to be cut in this position. The pBR322 fragment used seems to have been designed to include the ampicillin-resistance gene and *ori* but delete poison sequences in the plasmid which inhibit its replication in COS cells and I cannot identify any convenient restriction site that would allow a fragment of this nature to be obtained by an alternative route.

(ii) A linker of unknown sequence, containing a *Hin*dIII recognition site (light blue in Figure G, Annex 3), to be attached to the end of the pBR322 sequence at nucleotide 2448 to facilitate subsequent manipulations. As I have said the sequence of this linker is not given. Furthermore, the '605 patent does not provide any instructions on how to attach this linker to the end of the pBR322 sequence.

(iii) A 342 bp SV40 fragment comprising nucleotides 5171(via 5243 and then 1) through 270. I presume this to be a *Pvu*II-*Hin*dIII fragment, but some doubt is cast by the fact that *Pvu*II cuts the sequence to leave a blunt end at 272. This fragment contains the SV40 origin of replication and powerful late promoter.

(iv) A second fragment of SV40 DNA ostensibly comprising nucleotides 2553 through 2770 and containing a polyA-addition site. My conclusion is that '2553' is a mistake, since nucleotide 2553 of SV40 does not appear to be in or near a convenient restriction enzyme target site. I have assumed that the true number should be 2533, the position of a *Bam*HI site in SV40 DNA. 2770 is the position at which *Bcl*I cuts SV40 DNA. The '605 patent explains that this fragment was placed in the vector in the correct orientation 'via the unique *Bam*HI site'. This is an insufficient instruction, since the cohesive ends of the fragment, derived by *Bcl*I and *Bam*HI, are identical. An additional stratagem is required to ensure the correct orientation or to check the orientation of the inserted fragment in several isolates.

(v) A linker of unknown sequence (orange in Figure G, Annex 3), containing a *Bam*HI site to be attached between the two SV40 fragments.

(vi) Also present in the pDSVL1 vector is a 2,500 base-pair mouse DHFR 'minigene' isolated from a previously described plasmid, pMG-1 (Gasser *et al.*, 1982). The DHFR minigene of pMG-1 is excisable on an *Eco*RI-*Pst*I fragment. The patent provides no information on how the putative *Eco*RI-*Pst*I fragment from pMG-1 was ultimately ligated to a putative *Bcl*I cohesive end from the SV40 fragment containing the polyA-addition site in pDSVL1.

98. The sundry components of pDSVL1, the DHFR minigene (*Pst*I-*Eco*RI), pBR322 *ori* plus ampicillin-resistance determinant (*Eco*RI-*Hin*dIII linker), the SV40 *ori/pL* (*Hin*dIII-*Bam*HI linker) and the SV40 polyA-addition site (*Bam*HI-*Bcl*I) should ligate to each other, end-to-end in a predetermined order, apart from the ambiguity about the *Bcl*I-*Bam*HI orientation and the ligation of the DHFR minigene with the SV40 fragment containing the polyA-addition site. Some method of joining the *Pst*I end (DHFR) to the *Bcl*I end (SV40 fragment) would have to be designed to allow circularization to give plasmid pDSVL1. The unique *Bam*HI site provided in the linker between the two SV40 fragments forms the 'cloning site' for introduction of the EPO coding sequence on a *Bam*HI-generated fragment. The residual sequence of the linker, between the promoter and the inserted DNA, remains unknown.

99. Plasmid pDSVL1 was not deposited, so that the skilled person trying to follow the '605 patent must try to construct a version for his (her) own use. Because of the lack of specification of the linkers, the method of cutting pBR322 at nucleotide 2448 and attaching a *Hin*dIII linker to it, the lack of a stratagem to ensure the correct orientation of the *Bcl*I-*Bam*HI SV40 fragment and the method for including the DHFR minigene in the plasmid, the skilled person would be highly unlikely to be able to re-construct the identical plasmid.

100. Construction of a similar plasmid would have been facilitated by more detail and the avoidance of factual error in the descriptions. It would be possible to follow the general approach outlined in Example 6 to generate an expression construct using the same functional genetic elements as those described in the example. Such an exercise would involve a considerable amount of work for the skilled man, the extent of which would not be relieved by the teaching in the '605 patent.

Example 7A

101. This represents Amgen's first attempt to express the human EPO gene in COS-1 cells. The example is mainly directed to the creation of the expression vector used to transfect the COS-1 cells, pSVgHuEPO.

102. pSVgHuEPO comprises a plasmid, pSV4SEt and a genomic EPO fragment. Again, pSV4SEt was constructed "in-house" and is a plasmid that replicates both in *E.coli* and in COS-1 (monkey kidney) cells. Fig H(2), Annex 3, is an attempt to understand and explain the construction of pSV4SEt.

103. pSV4SEt was made from the following components :

(i) The fragment comprising nucleotides 2448 through 4362 of plasmid pBR322. This fragment contains the origin of replication (ori) and selectable marker gene (Amp^R) for *E. coli*. As I mentioned in the section on Example 6, although the fragment of pBR322 is defined, there is no teaching in the '605 patent nor is it obvious to me how this fragment was obtained;

(ii) A linker of unknown sequence (light blue in Figure H(2), Annex 3), containing a *Hin*dIII recognition sequence, to be attached to the end of the pBR322 sequence at nucleotide 2448 to facilitate subsequent manipulations as above. I refer to my comments at 97(ii) above;

(iii) An SV40 fragment comprising nucleotides 5171 through 270. This fragment contains the replication origin and promoters P_L and P_E of SV40. I presume this to be a *Hin*dIII to *Pvu*II fragment, but some doubt is cast because *Pvu*II cuts the sequence to leave a blunt end at 272;

(iv) A linker of unknown sequence (orange in Figure H(2), Annex 3) with an *Eco*RI recognition site inserted adjacent to the 270 end of the 5171-270 SV40 sequence;

(v) A linker of unknown sequence (pink in Figure H(2), Annex 3) with a *Sal*I recognition sequence inserted adjacent to the 5171 end of the 5171-270 SV40 sequence;

(vi) A fragment of SV40 DNA from nucleotides 1711 to 2772. This is probably a *Hin*dIII – *Bcl*I fragment, resulting from cutting at 1708 (*Hin*dIII) and 2773 (*Bcl*I).

If so, doubt remains about subsequent treatment of the fragment to manipulate the ends;

(vii) A linker of unknown sequence (yellow in Figure H(2), Annex 3) providing a *Sal*I recognition site inserted adjacent to the 2772 end of the 1711-2772 SV40 sequence.

104. As shown in Figure H(1), Annex 3, a 5.6 kb *Hin*dIII-*Bam*HI fragment encoding for EPO was excised from the recombinant phage lambda hE1 DNA and transferred into plasmid vector pUC8 creating pUC8-HuE, to generate a convenient source of the genomic EPO-coding sequence.

105. The EPO gene was then transferred from pUC8-HuE to a suitably digested pSV4SEt as a *Hin*dIII to *Bam*HI fragment (see figure H(3), Annex 3). Although this fragment contains the EPO-coding sequence, it was not established whether it contained a functional EPO promoter or, if it did, whether that promoter would work in transiently transfected COS-1 cells.

106. The lack of the sequence for the linker containing the *Hin*dIII sequence is of particular importance as this linker ends up immediately upstream of the EPO fragment in the expression plasmid, pSVgHuEPO (see Fig H(3), Annex 3). This linker sequence is inserted into the plasmid in a position proximal to the promoter region of the genomic EPO fragment and as such may interfere with the initiation of transcription from this region. Also, as it is unclear from Example 7A in the '605 patent whether or not the endogenous EPO promoter is active in COS cells, any initiation of transcription upstream of this region would transcribe this unknown sequence into the mRNA. This omission of the sequence of a linker, which is placed in such a potentially influential position for the outcome of the expressibility of the EPO gene, is a defect in the clarity of the instructions.

107. Neither pSV4SEt nor pSVgHuEPO were deposited meaning that the skilled person trying to follow the '605 patent must try to construct a version for his (her) own use. Due to the lack of specification of the linkers and the method of cutting pBR322 at nucleotide 2448, I believe that it would be extremely unlikely that the skilled person would be able to reconstruct these plasmids.

108. The skilled person would be able to construct an alternative expression vector with the same general characteristics as those of the pSVgHuEPO used in Example 7A. Since we are unsure about the mode of expression of the EPO gene in this example, the '605 patent gives no guidance about the key features of the construct that are required to ensure expression of the EPO gene. The skilled person would be left with essentially the same task as the patentee, if seeking to express the EPO gene.

Example 7B

109. This example relates to the creation of an expression vector for the expression of the human EPO gene in COS-1 cells using the well characterised SV40 'late' promoter to drive transcription. The expression vector, pSLVgHuEPO, is created by inserting a specially prepared EPO fragment into pDSVL1, the plasmid previously used in Example 6 (See Fig I, Annex 3, generally).

110. The EPO fragment for insertion into pDSVL1 was obtained as follows:

(i) pUC8-HuE (see Example 7A) is cleaved using *Bam*HI and *Bst*EII to obtain an EPO fragment with an intact coding sequence but with its promoter removed (See Fig I, page 1, Annex 3). Table VI of the '605 Patent gives the incorrect sequence for the *Bst*EII site at 43-49 nucleotides upstream of the EPO coding sequence. I shall discuss the significance of this below;

(ii) The resulting 4.9 kb *Bst*EII/*Bam*HI EPO fragment is then ligated, using a synthetic linker (turquoise in Figure I, Annex 3) of unknown sequence at the *Bst*EII end.

(iii) This fragment is then inserted into the plasmid pBR322 at the B*am*HI and *Sal*I sites forming the plasmid pBRgHE (see Fig I, page 1, Annex 3).

(iv) pBRgHE is then digested with *Bam*HI to produce an EPO fragment with *Bam*HI restriction sites at each end.

111. pDSVL1 is then digested with *Bam*HI and the EPO fragment inserted downstream of the SV40 late promoter and upstream of the SV40 poly A-addition site forming the expression vector pSLVgHuEPO (see Fig I, page 2, Annex 3).

112. Part of the synthetic linker sequence which was used to clone the *Bst*EII-*Bam*HI fragment containing the EPO gene and to generate an 'upstream' *Bam*HI cleavage-

site, lies between the SV40 promoter and the EPO coding sequence (see Fig I, page 2, Annex 3).

113. This linker sequence will be transcribed into the 5' untranslated region of the EPO transcript, a position in which it could affect mRNA secondary structure, splicing of pre-mRNA, or the initiation of translation. It is a serious defect in the clarity of the instructions to omit the sequence of a linker that is placed in such an important and potentially influential position for the outcome of the expressibility of the EPO gene.

114. As mentioned above, Table VI gives the incorrect sequence for the *Bst*EII site at 43-49 nucleotides upstream of the EPO coding sequence. This site is used for cleavage to isolate the EPO gene on a *Bst*EII-*Bam*HI fragment (see para 110(i) above). This fragment is placed downstream of the SV40 late promoter to form the basis of the expression construct for use in transformed CHO cells. The mistake in the sequence in the '605 patent, with GGTGACC instead of GGTCACC, would prompt the design of a linker sequence, to facilitate the ligation of the *Bst*EII end to a *Sal*I-generated end, that is incorrect and would give, at best, very inefficient ligation of the desired fragments (see Fig I, page 3, Annex 3). The mistake is particularly misleading because both the wrong and the correct sequences are authentic sites for recognition and cleavage by *Bst*EII, which cleaves the sequence GGTNACC, where N is any nucleotide.

115. Faced with failure to achieve the requisite ligation due to the misleading sequence, the skilled team would have to explore the reasons for failure, or devise another cloning strategy. There are viable options for each of these approaches, but they would inevitably involve the team in additional work that would have been avoided by accurate sequence information in Table VI. If lucky, the additional work may have merely involved troubleshooting the ligation reaction, though it is probable that this error would have required the sequencing of this region to obtain the correct sequence and then the redesign of the linker in order to proceed with the protocol.

116. Neither pDSVL1 nor pSLVgHuEPO were deposited meaning that the skilled person trying to follow the '605 patent must try to construct versions for his (her) own use. Due to the problems mentioned in my section on Example 6 relating to the construction of pDSVL1, the lack of information on the synthetic linker sequence and

the incorrect sequence given by Table VI for the B*st*EII site at 43-49 nucleotides, I do not believe the skilled person would have been able to reconstruct pSLVgHuEPO.

117. In conclusion, the skilled person would be left to design his/her own version of the expression construct for use in COS cells. The SV40 promoters and polyadenylation sites would be likely components, due to the contemporary practice, and the design and amount of work involved would not be reduced by the teaching of the '605 patent.

Example 10

118. The procedure in Example 10 involves, in the main, transfection of a DHFR$^-$ CHO cell-line with the expression vector for the EPO gene, pDSVL-gHuEPO containing a DHFR gene that allows selection for stable transformants. It appears that pDSVL-gHuEPO is the same expression vector as that used in Example 7B, i.e. pSLVgHuEPO.

119. Example 10 also involves the co-transfection of a DHFR$^-$ CHO cell-line with another expression vector for the EPO gene, pSVgHuEPO (see Example 7A), and a plasmid containing a DHFR gene that allows selection for stable expression.

120. As mentioned in my paragraphs on Examples 7A and 7B above the expression vectors, pDSVL-gHuEPO and pSVgHuEPO, were not deposited meaning that the skilled person trying to follow the '605 patent must try to construct versions for his (her) own use. Due to the reasons I mention above, I do not believe that the skilled person would have been able to reconstruct these expression vectors. My comments in the above sections on Examples 6, 7A and 7B, on the amount of work required for the skilled person to rework those examples, also apply to the molecular biology aspects of Example 10 as the same expression vectors are used.

STATEMENT

I understand my duty to the Court and have complied with that duty. I believe that the facts I have stated in this report are true and that the opinions I have expressed are correct.

SIGNED .....WJBrammar.....

PROFESSOR WILLIAM J BRAMMAR

DATED .....9th November 2000