SPECIAL EDITION
JANUARY 1986

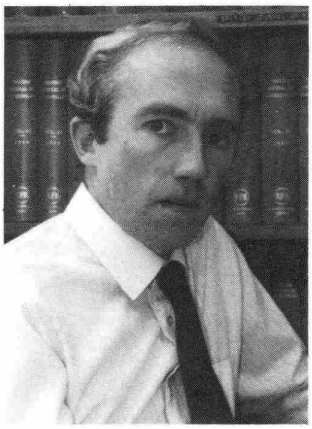
The Bulletin

Patron, H.R.H. The Duchess of Kent

Member of the World Federation of Hemophilia
Registered in accordance with the Charities Act 1960 (288260)

THE HAEMOPHILIA SOCIETY
P.O. Box 9
16 Trinity Street
London SE1 1DE
Telephone: 01-407 1010

GENETICALLY ENGINEERED FACTOR VIII



DR. E. G. D. TUDDENHAM
MD, MRCP, MRCPath
Senior Lecturer in Haematology
The Royal Free Hospital

A major goal of the biotechnology industry was reached early one morning in April 1984 when Gordon Vehar and Dan Eaton at the laboratories of Genentech in San Francisco detected factor VIII activity in culture fluid. The excitement they felt when factor VIII activity first appeared in a culture of hampster kidney cells was the culmination of 18 months work by a team of over 50 scientists. The human factor VIII gene had been completely cloned 3 months earlier but getting it expressed, that is translated from genetic information into active protein, factor VIII itself, proved even harder than expected. Almost a hundred different ways of splicing the gene into modified plasmids (see back) were tried and scores of different cell lines were infected by the plasmids carrying the human factor gene VIII before the right combination was hit upon. To clone (see back), the factor VIII gene, it had first been necessary to obtain some completely pure factor VIII in order to determine its structure.

This takes me back to the very beginning of the story. Factor VIII was first detected as long ago as 1911 by Addis in Edinburgh, who showed that addition of a small amount of extract from normal blood would correct the prolonged clotting of haemophilic blood. This observation was further refined over the years, notably by Taylor at Harvard. In the early 1950's a quantitative assay for a substance with the property of shortening the clotting time of haemophilic blood was devised in Oxford. This assay remains the primary method for detecting factor VIII but can go no further in characterizing the shadowy substance behind the clotting effect. A vast literature developed concerned with efforts to purify factor VIII out of blood and considerable confusion arose, because, as we now know, it is carried in the blood by another protein with quite distinct properties called, von Willebrand's factor. Various schemes for purifying factor VIII in fact concentrated mainly von Willebrand factor of which there is 50 times more per unit of blood than there is of factor VIII itself.

Adding to the difficulty is the fact that there is very little factor VIII indeed in blood and it is quite unstable, tending to be broken down by blood enzymes, so that, as many workers found to their frustration, the acitvity disappeared as one came closer to the goal of total purity.

My own efforts to purify this elusive factor began in 1976 when I was working with Leon Hoyer in Connecticut. We developed a partially successful system based on antibodies to von Willebrand factor which enabled us to obtain very small quantities of factor VIII free of the von Willebrand protein.

On returning to England to The Royal Free Hospital, I was able to continue this approach using antibodies and by 1982 we had devised a multistep procedure (see figure 1) that yielded small but useful quantities of totally pure factor. There were 2 key aspects of our process that enabled us to succeed where others had failed.

Firstly, through collaboration with David Heath and Sarah Middleton of Speywood Laboratories, we had access to their polyelectrolyte technology, this allowed very large scale partial purification removing most of the bulk contaminants.

Secondly, through collaboration with Alison Goodall in the Immunology department at The Royal Free, we began to raise monoclonal antibodies (see back).

The bulk processing capacity of polyelectrolyte combined with the exquisite specificity of monoclonal antibodies overcame all remaining problems, although Frances Rotblat and Don O'Brien, in my laboratory, found that they needed to use large amounts of highly poisonous nerve gas type enzyme inhibitors to keep the factor VIII stable.

In November 1982 I was able to demonstrate the purity of our product at an International Meeting on factor VIII in San Diego and shortly after that David Heath and I concluded an agreement with Genentech of San Francisco in a three-way partnership to clone and express through recombinant DNA technology synthetic human factor VIII (see back). The contribution of my laboratory was to continue purifying as much factor VIII as we possibly could and to supply it together with monoclonal antibodies and other reagents to the molecular biology and protein chemistry groups at Genentech. The agreements were finally signed just before Christmas and the target date for the first production was set at 18 months from commencement. Then began the race in which four other biotechnology groups were competing.

Cloning a gene for a particular protein presents a classical "needle in the haystack" type problem, for there are over one million structural genes occurring in a seemingly random fashion all over the immensely long coils of DNA which contain the coded genetic information.

In human cells there are 23 pairs of coiled DNA which are called chromosomes. These are large enough to be seen under the light microscope at certain stages of cell division and one

can then recognise that one pair of chromosomes, the sex determining pair, are similar in females but different in males. The odd one out is the Y chromosome which simply determines maleness. Its opposite partner, the X chromosome carries the gene for factor VIII as well as the gene for factor IX. In females there are of course 2 X chromosomes which accounts for the well known fact that female carriers of haemophilia A or B are usually not troubled with bleeding since they will have one normal gene to protect against the effect of a mutant haemophiliac gene on the other X chromosome.

Now a number of strategies have been evolved for fishing out individual genes. One approach that is very attractive and was used in the famous cloning of the insulin gene, is to identify a tissue or cell that is already synthesizing the protein whose gene you wish to clone. In that tissue or cell, the gene is already being transcribed from DNA in the chromosome into riboseneucleic acid RNA and the RNA is carrying the information to the ribosome which is the cell's device for turning genetic information into protein. This cuts down the number of different genes that one is sorting through from the total genetic information that is contained on the chromosomes to just those genes that are being expressed in the particular cell.

For instance, the pancreatic islet cells from which insulin was cloned, make a large amount of messenger RNA for insulin. Of course they also produce all the other housekeeping proteins, that are required to maintain the gloriously complicated process of life. So it is still quite difficult to find the insulin RNA.

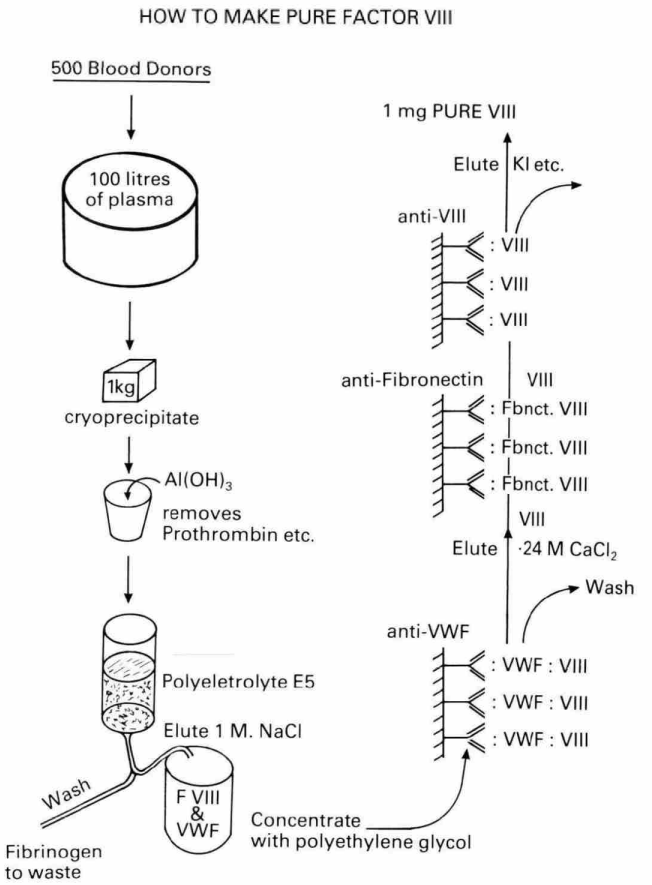
In the case of factor VIII, we were not at all sure which tissue was making factor VIII, or indeed, if any tissue was making enough factor VIII, for there to be retrievable quantities of RNA for factor VIII.

The approach which proved successful, is the one that I will now describe and attempt to give some idea of how extraordinarily labour intensive it was. The speed with which it was in fact accomplished is a tribute to the large numbers of highly dedicated people who worked on this project and to the exceptional level of coordination and cooperation that we achieved across an ocean and a continent.

The trick that was used, was to reverse the genetic code. Now the flow of information in the cell is from DNA of gene via RNA to protein. Frances Crick called this the central dogma of molecular biology and the way in which genetic information is translated into protein is as follows:-

DNA consists of long spirals of a sugar polymer (deoxyribose) on which are strung like pearls on a necklace four different kinds of molecules called nucleotides. These are adenine, thymine, guanine and cytosine or A, T, G and C for short. A given stretch of DNA might have the sequence of nucleotides G G G, C C C. If this were then transcribed into RNA, the specific base pairing of nucleotides would come into play. G always pairs to C and C to G, so a sequence in the RNA that corresponded would be C C C, G G G. The ribosome would then read this message in threes, according to the dictionary in figure 2. Thus two aminoacids would be added to a growing peptide chain and the two would be proline and glycine (see figure 3). If one obtained the protein and established that there was the amino acid sequence proline glycine, then one could translate back to infer the sequence of nucleotides in the gene. Here one would run into a problem since you will notice that more than one triplet codon is specific for glycine. To be precise there are 6 different triplets for glycine and 4 for proline. A more favourable aminoacid sequence for this kind of back translation exercise would be methionine, lysine, since methionine has only one codon and lysine only two.

The next step is to synthesise a short length of DNA with a sequence of nucleotides corresponding by the base pairing rules to the predicted sequence in the gene. Such a stretch is referred to as an oligonucleotide.



Fibrinogen, Fibronectin and Von Willebrand factor (VWF) were the main contaminating proteins which had to be removed, they are all literally very sticky which made them hard to wash away.

1 mg = One thousandth of a gram

Figure 1. *The multistage procedure by which altogether 20 mgs of pure factor VIII were obtained from about 2 metric tonnes plasma.*

In order to obtain aminoacid sequence from our purified factor VIII, it was first digested with the enzyme trypsin. Trypsin is one of the enzyme of normal digestion whose role is to cut up the proteins in our food into short lengths which can be easily absorbed. Factor VIII, it turned out, consists of a very long stretch of aminoacids indeed. In fact, over 2 thousand. Trypsin cuts it up into about 2 hundred short lengths of aminoacid called peptides. To separate these is difficult, but they can be partially resolved with the powerful modern analytical technique of high pressure liquid chromotography. (HPLC).

Figure 4 shows the detector readout from an HPLC run of trypsinized factor VIII, each little spike corresponds to a different peptide or mixture of peptides. The arrow is above a single peptide whose aminoacid sequence was the key to unlocking the whole factor VIII mystery. Sequencing peptides from the tiny amount available in this sort of experiment is another highly sophisticated area of technology. Rooms full of expensive complicated chemical apparatus programmed by computers are required to remove aminoacids one at a time and identify them precisely from amounts of peptide as low as a few millionths of a gram. The following sequence gave the key, (single letter aminoacid code, see table

I). A W A Y F S D V D L E K. Back translating an oligonucleotide probe 36 bases long was synthesised as follows:- C TTTTCCAGGTCAACGT CGGAGAAATAAGCCC AAGC. This is one of many possible choices but it turned out to be almost exactly the correct one thanks to known codon preferences, Dick Lawn's intuition and luck. This could then be used to probe for corresponding sequences in a library of human gene fragments. The human gene fragments were made by taking the genetic material from a cell line containing 4 X chromosomes not the usual 2. The DNA was cut into shorter lengths by means of enzymes which cut at specific nucleotide sequences. Such enzymes are called restriction endonucleases and are the critical tools which enable genetic manipulation. The shorter stretches of human DNA are then recombined with virus DNA. The viruses chosen for this purpose are able to multiply inside bacteria and are called phages. Because they multiply by faithfully recopying their own gene sequence along with that of any inserted sequence, one can obtain indefinitely large numbers of perfect copies of any desired DNA sequence. This is the basis of the famous Cohen Boyer patent on which the genetic engineering industry and indeed much of modern genetics is now based.

In practice, what is obtained, is a bank of different recombined DNA fragments covering the whole human gene sequence copied into phage (or in some cases, smaller replicating elements called plasmids), which are themselves carried in bacteria. The phage (or plasmid) is unable to replicate on its own since it must use the genetic machinery of a bacterium to do this. The gene bank then consists of thousands of bacteria each containing a different phage with a different stretch of inserted DNA. The bacteria can be spread thinly over the surface of a suitable growth medium where they will very rapidly multiply and divide. If spread thinly enough colonies develop, each one derived from a single different bacterium containing its own particular recombinant phage. The oligonucleotide corresponding to part of the protein sequence of factor VIII was made radioactive and used to probe for a complimentary DNA sequence in such a gene bank.

Figure 5 shows the localization of the radioactive probe over 2 bacterial colonies that contained the first fragments of the human factor VIII gene. (lower panel arrows). In the upper panel, in a different experiment the arrows point to where the same probe is binding 4 times more strongly (lane 2) to DNA from the 4X cell line than to DNA from a male 1X cell line (lane 1). The two independent but overlapping clones were sequenced and the sequence of their DNA corresponded in part precisely to the aminoacid sequence A W A Y F S D V D L E K.

The next part of the exercise was to use these clones to derive new probes that could be used to detect flanking overlapping complementary DNA clones within the same gene bank. This process is known as 'gene walking' a phrase which graphically illustrates the procedure. The initial clones covered only a small proportion of the whole factor VIII gene, about two thirds of the way along from the starting point and one attempted to walk in either direction up and down the gene until the whole gene was encompassed in overlapping clones. Figure 6 shows

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

Figure 2. *The genetic code dictionary. U C G and A are the nucleotides found in RNA. In DNA, U is substituted by T. Note that there is more than one nucleotide triplet corresponding to most aminoacids. Three triplets code for 'stop' which tells the protein systhesis machinery to stop translating. The code word for 'start' is that corresponding to methionine AUG so all proteins start with a methionine aminoacid. This may later be removed.*

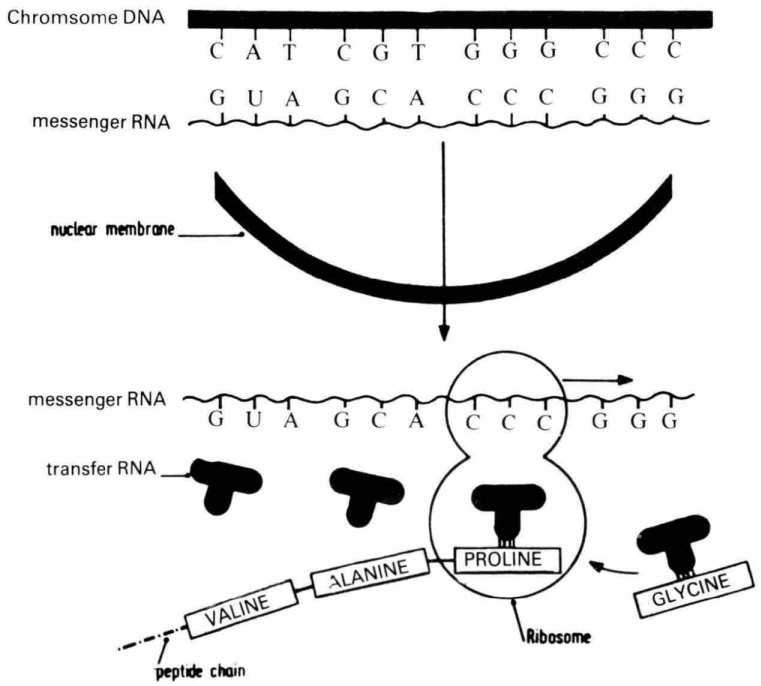


Figure 3. *The flow of genetic information from DNA via RNA to protein synthesis.*

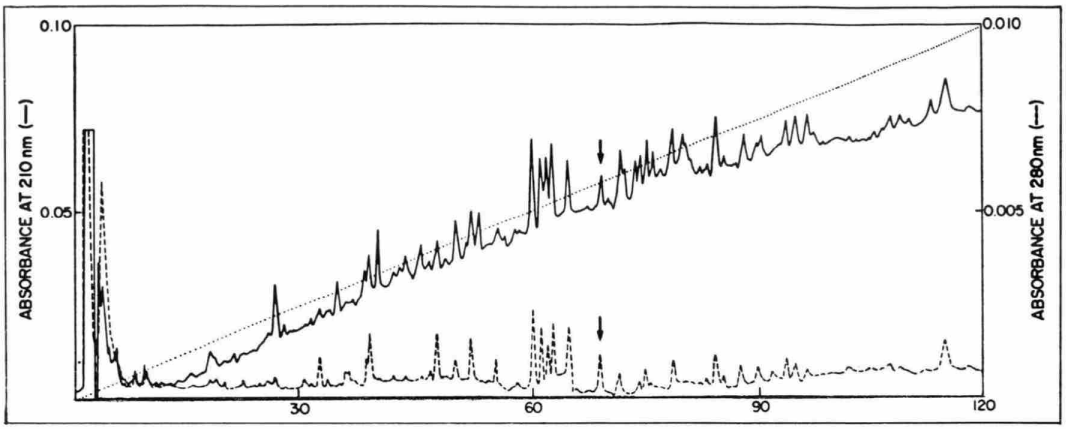


Figure 4. *Purifying the peptides of factor VIII. Each spike on the trace corresponds to one or more short length of aminoacids released from factor VIII by digesting it with trypsin. The arrow indicates the key peptide.*

this procedure in a diagramatic form together with a plan of the complete gene. The shaded overlapping bars in the bottom half of the figure represent the cloned DNA fragments that were copied in phage in the gene bank.

The first two clones were numbered 1 1 4 and 1 2 0. The next two overlapping clones which extended the cloned gene up and down stream respectively were numbered 222 and 482. After that in order to speed up the process a new gene bank was constructed that could contain much larger stretches of DNA spliced into a carrier called cosmid. The first new clone did not extend much further upstream, number P541, but P542 carried the cloning considerably further down stream. The phrase up and down stream refers to the usual direction of copying when genes are transcribed into RNA.

The middle section of the diagram shows points at which different restriction enzymes will cut the gene and this is information that was used in piecing the jigsaw together. The top lines of the diagram show firstly the scale in thousand of base pairs (KB). Next the structure of the gene itself. You will notice that there are shaded and unshaded regions. The unshaded regions are called introns, and do not code for aminoacid, in fact their function is unknown, but that portion is spliced out of RNA before it gets translated into protein. The dark shaded regions, which are called exons carry all the information for the aminoacid sequence for factor VIII. Altogether there are some 26 exons. The whole gene spans 185,000 base pairs and the exons 8,900 base pairs.

While one group of scientists headed by Richard Lawn were characterizing the human factor VIII gene, another group, under Bill Wood, had started to clone factor VIII from another direction, by copying RNA back into DNA. They were able to do this because the initial factor VIII gene clone could be used to test many tissues and cell lines for presence of messenger RNA for factor VIII. Quite fortuitously a cell line was found that had in fact been derived from the cells of a patient with a tumour of the lymphnodes. These cells were producing RNA for factor VIII, though not factor VIII itself. This RNA could be copied using an enzyme from a virus that reverses the usual flow back from RNA to DNA (the enzyme is called reverse transcriptase). A gene bank was then constructed containing reverse copy DNA from the RNA of this cell line and the clones were probed with a DNA stretch 189 bases long derived from the initial factor VIII clone 114.

This cloning exercise then proceeded largely independently of the genomic cloning which I described first. By sequencing DNA from both sources an independent check could be made on the accuracy and consistency of the results. Also, and critically for the purpose of the whole project, a full length copy of the messenger RNA coding for factor VIII protein could be assembled in order to attempt synthesis.

As I mentioned at the beginning this proved to be very difficult indeed but success was obtained with the expression vector (see back) shown in

Figure 5. *Locating the first cloned factor VIII gene fragments. A radioactive synthetic DNA sequence corresponding to the protein sequence of the peptide arrowed in figure 4 is shown here binding to two clones containing fragments of human gene copied in viruses that are growing in colonies of bacteria. These then appear as two dark dots arrowed. In the upper panel, the same probe binds 4 times more strongly to the genetic material from a cell line with 4 X chromosomes than to material for a male cell line with only one X chromosome (see text).*

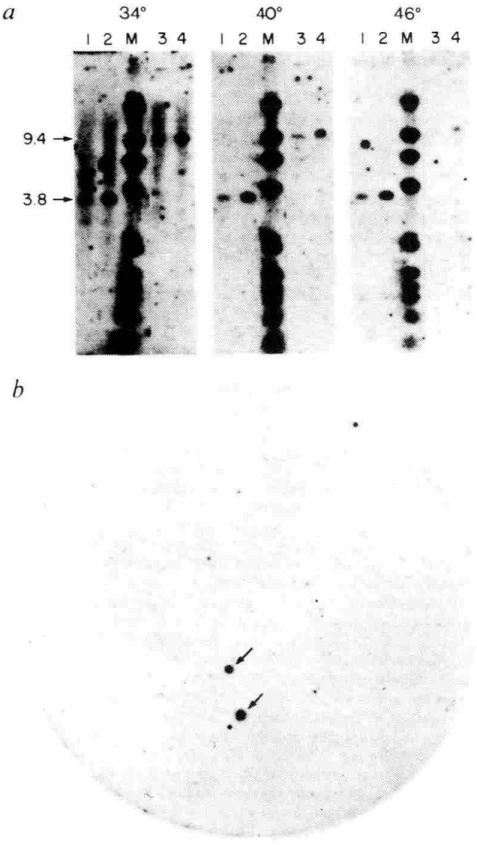


figure 8. This consists of a number of pieces of DNA taken from different organisms and viruses and put together in a way which should promote transcription of the factor VIII gene and synthesis of the protein.

Starting at twelve o'clock and proceeding clockwise, E and AML are promotor sequences (see back) taken from a monkey virus SV40 and a common cold virus the adenovirus respectively.

Next there is part of a plasmid confering resistance to the antibiotic ampicillin and then follows the full length factor VIII DNA sequence. After that comes part of the terminator region taken from hepatitis B virus and then another monkey virus promoter region.

Next, labelled DHFR, is a mouse gene for an enzyme which confers resistance to the anticancer drug methotrexate. The reason for including this gene is that one can then use methotrexate to select for cell lines in which the whole expression vector is functioning at a high efficiency. Copies of this curious chimeric recombinant DNA molecule were then introduced into various cell lines. A well known cell line that grows readily in culture vessels is derived originally from a baby hampster's kidney. These are called BHK cells and one such cell line that had the factor VIII expression vector introduced in it started to make factor VIII.

Figure 9 shows the formal proof that factor VIII activity authentically identical to the natural product was made by these cells. In the upper panel culture fluid with about one unit of factor VIII clotting activity per ml is poured over a column to which monoclonal antibody specific for human factor VIII has been linked. The clotting activity falls to zero as the factor VIII is taken up by the column. Subsequently, the factor VIII can be recovered by washing off with a salt solution.

In the lower panel, a somewhat similar experiment was performed, but instead of antibody on the column there was chemically bound von Willebrand factor. Once again the column takes up the factor VIII and it can subsequently be recovered by the salt wash.

Other experiments showed that the synthetic factor VIII reacted with other blood clotting factors exactly like the natural product and now most recently large enough quantities have been made to show that the chemical structure is identical to plasma derived factor VIII.

The production process is being scaled up and already trials have been started with the material in haemophiliac dogs. At the present rate of progress, we are now predicting clinical trials for late this year (1986). This work is being carried out jointly with Cutter Laboratories who will manufacture and market the product worldwide.

Before the product is administered to humans it will of course have been rigorously tested for safety and purified to a degree thousands of times higher than any product a haemophiliac has received to date. In fact it will be necessary to add back to the purified factor VIII some protein to act as a carrier, such as albumin. The new product will be completely virus safe and no doubt fully active in regard to clotting. It will be necessary to test it for a number of years on a selected group of patients to be absolutely sure that there are no unanticipated side effects. If all goes well, we would expect synthetic factor VIII to replace the natural product fairly rapidly during the early years of the next decade.

There have been several very interesting spinoffs from this massive research project. Using the cloned normal factor VIII gene comparison could be made with the factor VIII gene from haemophiliacs. Most of the cases studied to date show no obvious abnormality in their factor VIII gene which indicates that the mutation is very localized within this huge stretch of DNA. There is at present no general method for locating single substitutions of nucleotides within a large region of DNA, but in some instances mutations will change the site at which a restriction enzyme would cut in the normal gene.

Jane Gitschier used DNA from some of our patients to study their pattern of gene fragmentation. Three mutations causing haemophilia have now been localised because they cause a change in restriction enzyme sensitivity. In each case the substitution of a single cytosine with a thymidine residue changes a codon that should read C G A for argenine to T G A which indicates stop or end translation. It may seem unfair that faulty punctuation can cause a disease as serious as haemophilia A but that is the way the genes work and very minor errors can have disastrous consequences.

The normal mechanism for copying DNA is wonderfully faithful and has an error rate far below that of any manmade machine and indeed most such tiny errors if they occur away from the exon regions of structural genes have no effect or merely lead to the minor variations which make us all different.

Again, using restriction enzymes, several minor variations in sequence have been found that are common amongst normal people in or close to their factor VIII gene. These minor variations can be detected using a factor VIII gene probe and are the basis for the new methods of carrier detection which are capable of giving a definitive yes or no answer to the woman who wishes to know if she is a carrier. Clearly this is a great advance on the coagulation factor methods we have had to use up to now which could never give a definite yes or definite no answer. A subsequent edition of The Bulletin will give further details of this exciting advance in genetic counselling that is now becoming available at a number of centres in the United Kingdom.

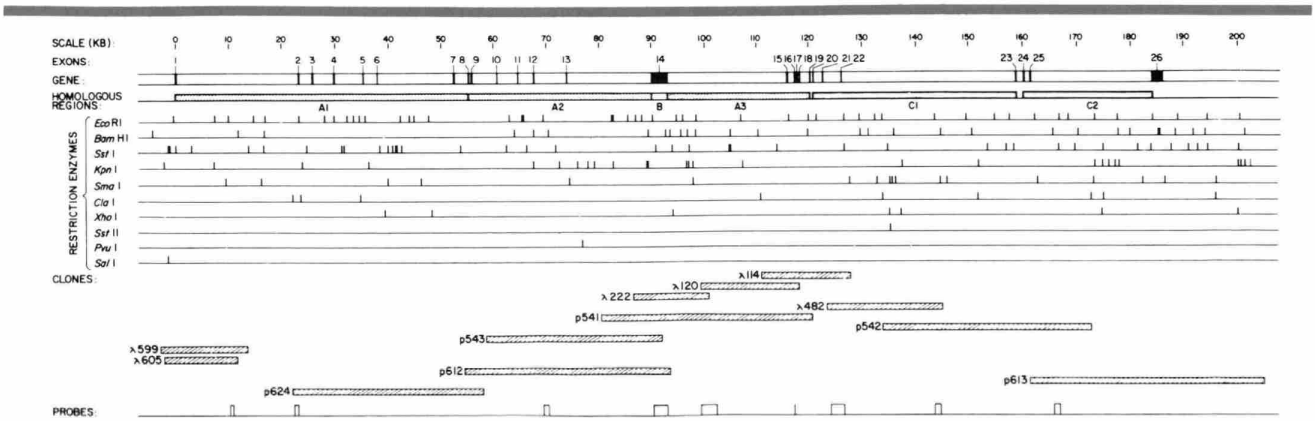


Figure 6. *Map of the human factor VIII gene. See text for detailed explanation.*

Finally using the factor VIII gene as a probe, we have been able to pin down once and for all where factor VIII is normally made in the body. As has long been suspected from a variety of experiments, the liver and spleen were found to contain the most amounts of messenger RNA for factor VIII. Within the liver, the hepatocyte turned out to be the cell producing factor VIII. Quite surprisingly we found that lymph node contained specific factor VIII RNA in large amounts and there were lesser amounts in other tissues, such as muscle, kidney and pancreas. The bone marrow was completely negative, which brings me to my final point. What are the hopes of gene therapy to cure haemophilia by reinserting a normal factor VIII gene into a sufferers body cells? A great deal of research is currently focused on gene therapy, particularly in relation to certain metabolic disorders and congenital anaemias. Bone marrow derived cells are seen as the most hopeful route for the reinsertion of active genes, since these cells are readily accessible for culture, manipulation and reinfusion. Absence of functioning factor VIII synthesis in these cells is rather against using them for gene therapy in haemophilia. Undoubtedly research will continue directed at methods of targeting genes to cell types and to particular locations within the chromosomes; there may be an issue of The Bulletin for the year 2001 which will describe the first cases of haemophilia cured by gene therapy. I certainly hope so.

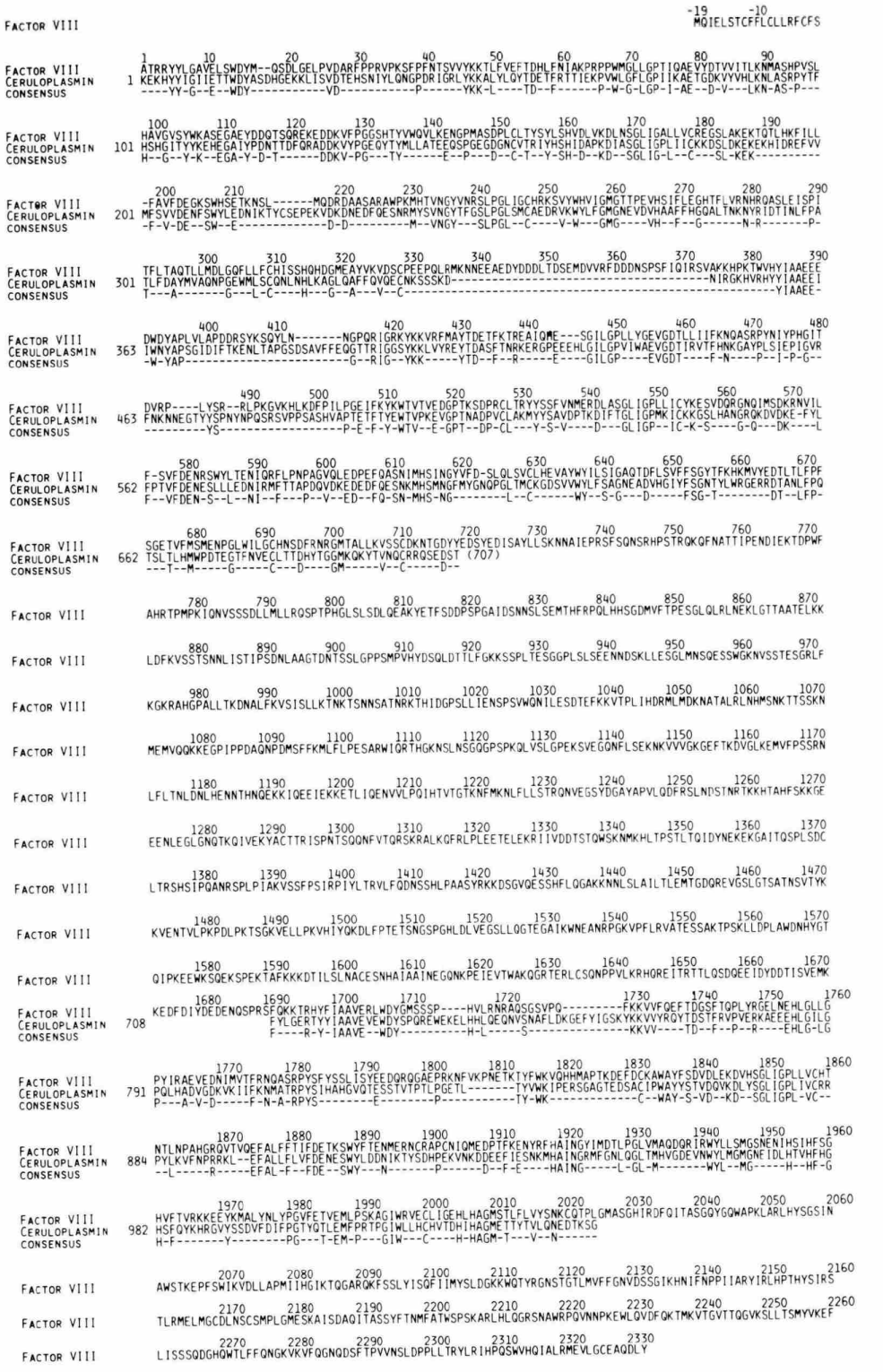


Figure 7. *The complete aminoacid sequence of human factor VIII, the single letter notation for aminoacids is used. The sequence proved to be very similar to that of a completely different protein, ceruloplasmin which is the main copper carrying protein in plasma. This means that the two proteins are descended from a common ancestral gene that probably diverged into two lines of descent at least a hundred million years ago. The protein starts with M for methionine but this and the next 18 aminoacids are removed during the secretion of factor VIII out of cells into plasma. See if you can find the AWAY peptide. Clue, it is about three quarters of the way along. Factor VIII has the longest single chain of aminoacids of any protein in plasma.*

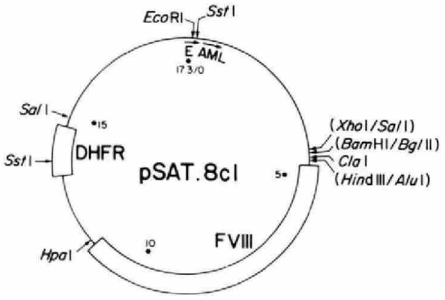


Figure 8. *Assembly of a full length factor VIII expression plasmid. When hampster kidney cells growing in tissue culture were infected with copies of this circle of recombined DNA from various origins (see text), they started to make human factor VIII.*

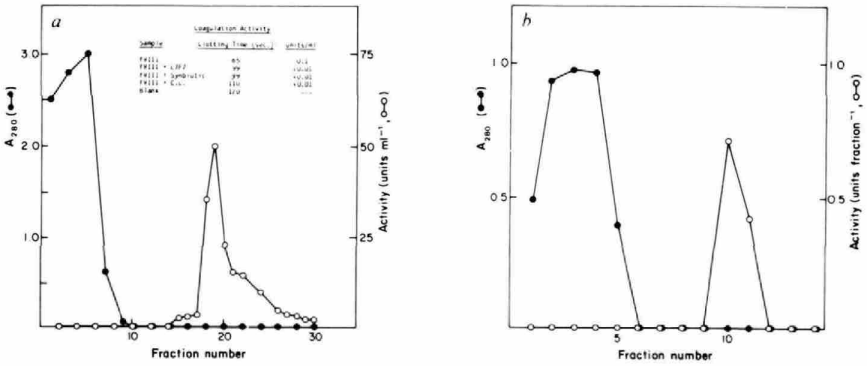


Figure 9. *Proof that culture fluid from hampster kidney cells infected with the factor VIII expression plasmid are producing authentic factor VIII. Left panel: the factor VIII binds to a specific antibody. Right panel: the factor VIII binds to von Willebrand factor.*

Figures 4-9 are reproduced by permission of the Editor of Nature, where they first appeared (references 2, 3 & 4).

— K	! A	W	A	Y	F	S	D	V	D	L	E	K ! —
	CGA	ACC	CGA	ATA	AAG	AGG	CTG	CAA	CTG	GAC	CTT	TTC Synthetic probe
	GCC	TGG	GCT	TAT	TTC	TCT	GAT	GTT	GAC	CTG	GAA	AAA Actual gene
	--x	---	---	---	---	--x	--x	---	---	---	---	--x

Trypsin ! — Trypsin !

--- Correct x incorrect

Figure 10. *Comparison of the protein, synthetic oligonu cleotide probe and actual gene sequences. An almost perfect match of probe and gene enabled recovery of part of the gene from a gene 'library'.*

REFERENCES

1 Purification of human factor VIII:C and its characterisation by western blotting using monoclonal antibodies.
Rotblatt, F., O'Brien, D.P., O''Brien, F.J., Goodall, A.H., & Tuddenham, E.G.D.
Biochemistry, 24: 4294-4300 (1985).

2 Characterization of the human factor VIII gene.
Gitschier J, Wood, W.I., Goralka T.M., Wion, K.L., Chen, E.Y., Eaton, D.H., Vehar, V.A., Capon, D., Lawn, R.M.
Nature 312, 326-330. (1984).

3 Expression of active human factor VIII from recombinant DNA clones.
Wood, W.I., Capon, D.J., Simonsen, C.C., Eaton, D.L., Gitschier, J., Keyt, B., Seeburg, PH., Smith, D.H., Hollingshead, P., Wion, K.L., Delwort, E., Tuddenham, E.G.D., Vehar, G.A., & Lawn, R.M.
Nature, 312, 3300-337 (1984).

4 Structure of human factor VIII.
Vehar, G.A., Keyt, B., Eaton, D., Rodriquez, H., O'Brien, D.P., Rotblat, F., Oppermann, H., Keck, R., Wood, W.I., Harkins, R.N., Tuddenham, E.G.D., Lawn R.M., & Capon, D.J.
Nature, 312, 337-343 (1984).

5 Genetic mapping and diagnosis of haemophilia A achieved through a Bcl I polymorphism in the factor VIII gene.
Gitschier, J., Drayna, D., Tuddenham, E.G.D., White, R.L. & Lawn, R.
Nature, Vol 314, No 6013, 738-740 (1985).

6 Detection and sequence of mutations in the factor VIII gene of haemophiliacs.
Gitschier, J., Wood, W.I., Tuddenham, E.G.D., Shuman, M.A., Goralka, T.M., Chen, E.Y. & Lawn, R.M.
Nature, 315, 427-430 (1985).

7 Antenatal diagnosis and carrier detection of haemophilia A using a factor VIII gene probe.
Gitschier, J., Lawn, R.M., Rotblat, F., Goldman, E., & Tuddenham, E.G.D.
Lancet,i: 1093 (1985).

8 Distribution of factor VIII m RNA and antigen in human liver and other tissues.
Wion, K.L., Kelly, D., Summerfield, J.A., Tuddenham, E.G.D. & Lawn, R.M.
Nature 317, 726-729 (1985).

(Dr. Tuddenham is now looking for a new project!)

Glossary of terms

Plasmid: circular piece of DNA (see below) that reproduces itself autonomously inside another cell.

Clone: all the cells derived from a single cell by repeated divisions and all having the same genetic constitutions. By extension this is also applied to DNA and then means identical copies of a defined DNA fragment.

Gene: a segment of the DNA molecule containing the information for synthesis of a specific protein or peptide.

Protein: a long series of aminoacids linked together by peptide bonds. Examples: keratin in skin and hair; albumin in plasma; Factor VIII, Factor IX, etc.

Peptide: a short series of aminoacids.

Aminoacid: a molecule containing an amino group and a carboxyllic acid group capable of forming peptide bonds and thus long chains of aminoacids called proteins. there are 20 different commonly occurring aminoacids in natural proteins. see Table I.

Chromosome: bodies situated within cells of higher organisms inside the inner compartment called the nucleus and composed of very long stretches of DNA together with protein around which the DNA is coiled.

DNA: deoxyribose nucleic acid, the chemical which contains the information that is passed on from generation to generation and which specifies the processes of life by directing synthesis of proteins. It exists in a complimentary paired structure that enables faithful copying and reduplication as cells divide and when egg cells and spermatazoa are formed. The genetic information is coded in DNA as the order of 4 different nucleotides C T A and G. Three nucleotides make up one codon and there are therefore 4 x 4 x 4 = 64 different codons. These specify the 20 different aminoacids as well as start and stop for protein synthesis. More complicated combinations specify many other processes which are only beginning to be understood, such as when and how much of a protein to make, and how the body is to be organised.

RNA: ribose nucleic acid. A single strand of nucleotides C U A G (U replaces T of DNA) which carries information from the chromosome to the protein synthesis machinery in a cell (the ribosome). This is messenger RNA and a different messenger is transcribed from each gene according the pairing A:U, T:A, G:C, C:G. Specialised RNA molecules also transfer aminoacids to the end of a growing peptide chain. Transfer RNA reads the genetic code on the messenger RNA.

Restriction enzyme: proteins found in bacteria which cut DNA at precisely defined sequences of nucleotides. eg Taq 1 cuts only at the sequence TCGA.

Recombinant DNA: DNA which has been recombined by manipulation in the laboratory into sequences that do not occur in nature.

Expression plasmid: A form of recombinant DNA designed to induce synthesis of a protein such as insulin, by cells which would not normally produce that protein eg bacteria.

Monoclonal antibody: antibodies produced in tissue culture by fusing cells from an immunised animal with tumour cells. The resulting hybrid cells then grow continuously and produce indefinitely large amounts of absolutely specific antibody.

Antibody: proteins produced by the body in response to challenge by foreign protein such as is present in bacteria and viruses and which bind specifically to foreign protein in order to hasten its removal. Antibodies are used in biochemistry laboratories as highly specific reagents to detect particular peptides or proteins.

Promoter sequence: A sequence recognised by the enzyme which starts RNA transcription from the DNA template. eg TATA.

TABLE I

AMINOACIDS Building blocks of protein

Fullname	Notation: Three letter	Notation: One letter
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y