The Legacy of Pehr Edman

Council of ASBMB has decided to introduce a new award for Post Doctoral Biochemists and Molecular Biologists as part of its series of Medals and Awards of the Society. The new award is named in honour of the late Pehr Edman, a distinguished protein biochemist who did much of his seminal work on protein sequencing in Melbourne. Lindsay Sparrow outlines his achievements and their impact on protein science.

Biologists have an extraordinarily valuable resource in the huge number of protein sequences now available in various databases and this resource is rapidly growing in value as new sequences are added. Most new entries in the protein sequence databases are now derived from translation of nucleic acid sequences but this was not always the case; until around the early 1980s, most proteins were sequenced directly by chemical methods. Early protein sequencing was an extremely laborious affair and it was not until chemistry for the sequential removal of amino acid residues from the N-terminus of peptides and proteins became available in the mid 1950s that sequences started to appear in the literature in any number. This chemistry was essentially developed by one man, Pehr Victor Edman.

Pehr Edman was born in 1916 and lived his early life in Stockholm where he was educated, eventually being granted the degree of Doctor of Medicine in 1946. His interest lay in research and his first independent work was carried out in the Rockefeller Institute during a one-year post-doctoral appointment in 1946-47. In 1947 he returned to a teaching position at Lund where he continued his research on the structure of proteins and started to make a substantial reputation for his work. In 1957 he moved to Melbourne as the first Director of Research at the newly established School of Medical Research at St Vincent’s Hospital. This position he held for 15 years before returning to Europe as the Director of Protein Chemistry at the Max Planck Institute at Martinsreid in 1972. With his premature death in 1977, the scientific world lost a major talent and Edman himself undoubtedly missed out on the full recognition that he deserved.

Early studies on protein sequencing

Pehr Edman’s research career focused almost exclusively on the development of methodology for determining the amino acid sequence of proteins. His first publication on the subject, in 1949, appeared before the insulin sequence had been completed by Fred Sanger and co-workers and it could be argued that only after the publication of the Sanger papers, in the mid-1950s, was the peptide nature of proteins conclusively proved and widely accepted. Edman’s 1949 paper was the first description of the use of the phenylisothiocyanate (PITC) reaction in the sequencing of proteins and this reaction is now known as the Edman reaction. Edman, building on earlier work of Abderhalden and Brockman and others using phenylisocyanate, recognised that replacement of the oxygen of the cyanate moiety with the more nucleophilic sulphur atom should yield a cleavage step that would proceed under relatively mild conditions. This was the key to obtaining high repetitive yields and thus to long sequence runs.

Clearly other chemists were thinking along the same lines at around this time as shown by publications that appeared from a number of laboratories; these publications described alternative reagents for the sequential removal of amino acid residues from the N-terminus of a peptide or protein. The chemistries used in these reaction sequences are similar to that of the Edman reaction, at least to the extent that they all give a thiazolinone as the product of an acid-catalysed cleavage step. However, none of these alternative reaction sequences seem to have been developed further and the protein sequencing community has not adopted them. This is probably due to the determination with which Edman pursued the development of the PITC reaction. Over the years 1950-56 Edman refined the conditions for the three reactions of coupling, cleavage and conversion that comprise the Edman degradation and described them in a series of eight important papers.

Until the early 1950s the principal method used for N-terminal analysis of proteins was the DNP method developed by Sanger and co-workers; in fact, the insulin sequence was determined using this technique and the few peptide sequence determinations published up to this time all used it. The impact of Edman’s work can be seen from the fact that almost immediately after publication in 1950 of his first major paper describing the PITC technique, it was taken up by other groups so that by 1952 Edman-derived sequences were starting to appear in the literature. A roll call of authors of sequence papers incorporating Edman chemistry at around this time includes, among others, such names as Choh Hao Li, Vincent du Vigneaud, Caesar Milstein and Brian Hartley. By 1960, even Fred...
Sanger himself was using the technique. An Australian connection is established by articles in which E.O.P. (Ted) Thompson and D.C. (Dennis) Shaw feature as authors, although admittedly while still working overseas. A number of variants of the Edman PITC technique were developed relatively quickly in various laboratories. The technique as described by Edman, had the N-terminal amino acid being identified directly as the PTH-derivative using TLC. Other widely used methods for identification of this residue were: the dansyl-Edman approach of Gray and Hartley, where the Edman chemistry is simply used to remove the N-terminal amino acid and expose the next residue for identification as the dansyl derivative; the subtractive-Edman technique where the N-terminal amino acid was identified as a difference between amino acid analyses before and after one cycle of Edman chemistry and back hydrolysis of the PTH-amino acid to the amino acid that was then identified by amino acid analysis.

Development of an effective method for determination of N-terminal protein sequences immediately made it possible to contemplate tackling the sequences of longer proteins. Work on some readily available proteins such as cytochrome c, ribonuclease, haemoglobin and lysozyme, to name a few, had been underway since the mid-1950s using the “classical” Sanger DNP technique. However, progress was very slow and a scan of papers from this period shows that it was very hard work indeed to even put four or five residues in sequence. The introduction of Edman chemistry led to dramatic improvement in the rate of accumulation of sequence data. Nonetheless, it is interesting to see that many laboratories did not immediately abandon the DNP technique but retained it for determination of N-terminal amino acids along with the use of Edman chemistry for actual sequencing. The sequences of ribonuclease (Hirs, Moore & Stein) and the coat protein of TMV (almost simultaneously by groups from Berkeley and Tubingen) were published in 1960.

By 1965 the number of known protein sequences was so great that it was necessary to establish the early sequence databases; among the first of these was the Atlas of Protein Sequence and Structure (edited by Margaret Dayhoff). By 1971 the Atlas contained 700 sequence entries, the majority of which had been determined using one version or other of the Edman degradation.

**The move to automated sequencing of proteins**

Having completed the development of the manual three-stage PITC degradation of peptides to his satisfaction, Edman became convinced of the need to have the chemistry automated. He felt that this would overcome some of the solubility problems associated with sequencing of proteins rather than peptides and would allow longer sequence runs to be carried out. By this time, 1961, he was firmly established in his Australian laboratory and had a research associate who was a master technician, Geoff Begg. Within two years they had solved the main problems in converting the coupling and cleavage reactions of the three step reaction sequence from a fully manual to an automatic process, leaving just the conversion step to be carried out manually in batches. Probably the most important single innovation incorporated in the new instrument was the glass reaction vessel in the form of a spinning cylindrical cup; the protein was applied as a thin film to the inner surface of this cup where it was readily accessible to the reagents used in the reactions. His success in producing an automated instrument was reported by Edman at a meeting in Scotland in 1964 but, typically, he continued with fine tuning for a further two years before full publication. The details of the Protein Sequenator were described in a paper in Volume 1 of the European Journal of Biochemistry in 1967 under the simple title “A Protein Sequenator” and with just two authors: Edman and Begg. This paper not only described the instrument in great detail but also presented the results of a sequence run – 60 cycles of N-terminal sequence on 5.0 mg (~0.3 µmoles) of humpback whale myoglobin – that was an outstanding achievement.
At Edman’s insistence and with the agreement of the board of the St Vincent’s School of Medical Research, the technology involved in the Sequenator was not patented. By 1973, there were over 100 instruments operating around the world based on the design of Edman and Begg, as documented by Hugh Niall in a comprehensive article in Methods in Enzymology. The most successful of the commercial spinning cup machines was the 890C, manufactured by Beckman Spinco under the guidance of Niall, an Australian former student of Edman’s. At least four other companies in the USA also made instruments, along with a Japanese company and another in Europe. A number of “home-made” Sequenators were put together in other laboratories, in some cases introducing modifications based on their own experience.

An alternative approach to protein sequencing was the development by Richard Laursen, John Bridgen and others in the late 1960s and early 1970s of the solid-phase Edman degradation. Here the peptide was attached by its C-terminus to a polystyrene resin and the sequencing chemistry was then carried out on the resin-bound peptide in a small column. This proved to be particularly effective for relatively short peptides, which tended to wash out of the spinning cup during the sequencing reactions. The reaction sequence was readily automated and solid-phase instruments using this principle were produced commercially.

A novel application of solid-phase Edman chemistry was that of George Stark who devised an insoluble isothiocyanate to which the peptide became attached in the coupling step and was released during the acid cleavage step of the sequencing cycle. This technology was brought back to Australia by Lyndsay Dowling of the CSIRO Division of Protein Chemistry (DPC) after a six-month study leave in the Stanford laboratory of Stark, but never matured into a generally used sequencing system.

Edman, himself, produced a second advanced Sequenator in collaboration with scientists at the DPC; in fact two instruments were assembled in parallel, one for the St Vincent’s laboratory and one for CSIRO. This exercise used the expertise of Edman and Begg and the excellent electronic and machine shop facilities of the CSIRO Division and was coordinated from the CSIRO point of view by Adam Inglis. Inglis also worked with Edman on other aspects of protein chemistry and together they produced a series of publications on the CNBr degradation of proteins, examining mechanistic aspects. The DPC went on to make another, even more advanced, Sequenator with a fully microprocessor controlled programming unit and for a time in the late 1970s and early 1980s ran two instruments in parallel.

Further developments of the spinning cup sequencer and of the technique for manual sequencing were undertaken by a number of laboratories through the 1970s. Notable improvements were the automation of the conversion step by Wittman-Liebold and the development of HPLC techniques for the identification and quantitation of PTH-amino acids. The use of reversed-phase chromatography in particular, led to a significant increase in sensitivity of both automatic and manual sequencing so that by 1980 Hunkapiller and Hood at Caltech were able to demonstrate the ability to sequence, in a modified spinning cup sequencer, at subnanomole levels of protein.

An automatic sample loader on an HPLC committed to PTH-AA analysis enormously increased the throughput of laboratories that used manual Edman sequencing. Indeed, in our laboratory in CSIRO’s DPC, during the mid 1970s to the mid 1980s when manual protein sequencing was a major activity, a skilled operator could achieve six cycles of Edman chemistry a day on each of eight peptides for a total of nearly 50 amino acid residues positioned in the protein chain. This was more than three times the rate of sequencing of a spinning cup sequencer and for this reason many laboratories did not immediately abandon manual Edman sequencing after acquisition of an automatic instrument.

The strategy for the chemical sequencing of proteins has evolved along with the increasing efficiency of the Edman degradation since its introduction in the mid 1950s. The early sequences published were those of relatively abundant proteins, often commercially available. These sequences were obtained through fragmentation of the protein into peptides that were relatively small and thus easily sequenced; such peptides were often separated and purified by high voltage paper electrophoresis. With the development of the automatic Protein Sequenator and its capability for long sequence runs the approach changed to one of obtaining a smaller number of longer peptides by using increasingly selective cleavage reactions. The demands of this strategy probably explain Edman’s interest in the CNBr cleavage of proteins, one of the most successful reactions for producing large protein fragments.
Further developments in protein sequencing

The next major progression in the use of the PITC reaction for protein sequence determination occurred after Edman’s death with the production of the gas-phase sequencer at Caltech in 1981. In this instrument the protein was subjected to Edman chemistry not spread on the inner surface of a spinning cup but rather adsorbed to a glass fibre disc in a small reaction chamber. The early versions of this instrument introduced the Edman reagents in the gas phase but subsequent models used a “pulsed liquid” mode of delivery. Further development and commercialisation of the gas phase sequencer was undertaken by the Applied Biosystems (ABI) company that had been formed specifically to exploit the technologies developed in Leroy Hood’s laboratory at Caltech.

By miniaturisation of all components, ABI were able to achieve sequencing down to the low picomole level of protein for the earliest model of the gas phase sequencer produced – the ABI 474. The latest model, the Procise, offers subpicomole sequencing. Hewlett Packard (now Agilent) have also produced a protein sequencer that does not use the spinning cup principle; in their instrument, the HP1000, the protein is bound to a reversed phase adsorbent, approximately 30µl in volume in a small column, and this effectively is the reaction chamber. The HP1000 also has the capability of low to subpicomole sequencing.

As the sensitivity of protein sequencing increased, more and more proteins, including low abundance proteins with important biological activities, became accessible to the technology. For example, in 1980 Hunkapiller and Hood described in *Science* the improved spinning-cup Sequenator with subnanomole capabilities. This paper was followed by a series of three short papers documenting the determination of around 20 residues of N-terminal sequence on each of three interferons isolated in subnanomole amounts from natural sources. At that time it would not have been possible to determine the complete amino acid sequence of any of these three proteins on the amounts available. However, from the mid-1970s onwards, developments in nucleic acid biochemistry increasingly influenced the approach to protein sequence determination.

Often only six to eight residues of amino acid sequence were needed to design a DNA primer that could be used to isolate and clone the gene (or more often, the cDNA) of a protein of interest. Obtaining the DNA sequence specifying the protein was then a relatively straightforward task and then, of course, the protein sequence was there for the translating. Thus the 20 residues of N-terminal sequence obtained for the three interferons discussed above ought to have been enough to allow cloning of their genes and thus determination of their entire sequences. For low abundance proteins this indirect approach is generally much quicker and easier than direct sequence determination via the protein route – indeed these days, it is almost the only technique used for obtaining the sequence of a protein of interest.
Professor Pehr Edman: A Personal Reflection

Elizabeth Minasian was a research assistant who worked closely with Edman. She provides an appreciation of his personality, his methods of working and his talents as a research biochemist.

I was fortunate enough to work in Dr Pehr Edman’s laboratory at St Vincent’s Hospital of Medical Research as a research assistant from beginning of 1962 to 1970 when the sequencing work was having its most important impact in Biochemistry. I have many fond memories of Edman, and I will highlight some of them that I think are important to illustrate his personality as a scientist and a man.

Pehr Edman’s name is renowned all over the world owing to his method of isothiocyanate degradation published in Archives of Biochemistry and Biophysics 22, 475 (1949) and Acta Chemica Scandinavica 4, 277 and 283 (1950). Based on this reaction, Edman and Begg decided around 1961 to build an instrument to automate the above procedure. I came into the field when the instrument was in its infancy, when everything was working manually. It took a good deal of subsequent work to make its operation systematic and build one of the basic tools of modern chemistry, the Sequanator.

Pehr Edman was accustomed to working alone. He produced his own problems, his own solutions and his own techniques. He was a man of precision and originality and he disliked being dependent on somebody else’s idea and experimental skills, unless he was confident of the precision of that person’s work. As a result of his lonely habits, Pehr Edman did not have many friends and was disliked by many colleagues.

He was a very disciplined and severe person, which created problems with most of his staff. He wanted his assistants to be very much interested in their work and to be skilled people. But indirectly, by the impact of his work, he had a profound influence in Biochemistry. As a research worker, Pehr was an experimenter in the fundamental developments in the chemistry of proteins. His investigations all lay within the sphere of preparative and engineering chemistry. Supported by his exceptional memory, habits and organising ability, he was able to master a great amount of experimental material.

Above all, Pehr Edman was a scientist, the father of modern Biochemistry. His methods enable us to study and understand many of the problems related to protein today. He considered even the most ordinary events from a serious scientific point of view and this attitude, combined with his peculiar ponderosity, which marked both his mind and his appearance, resulted for instance
in his being unable to appreciate a joke: he took everything seriously. He seemed to exude this scientific seriousness always and everywhere, but in fact it was only a shell protecting his real self. He believed that it would not be suitable for a scientist to be cheerful, except on very special occasions.

It was also rare for him to praise anyone’s work. He assumed a person would be further urged on to work harder next time, if he were not praised for his previous work. I remember him once coming to me just before he was leaving work and with a smile on his face, congratulated me on the precise manual degradation and identification of the sequence of the five steps which I performed on several normal and abnormal immunoglobulins. He checked all my results with the automated Sequenator and they were identical. The congratulation was directed to the skill which I developed in dissolving the insoluble PTC in the coupling buffer. He likened my accomplishment to that of a witch, which I did not quite appreciate at the time. However, I think he meant it as praise!

Another time I remember one day when Pehr Edman came to me and asked first thing in the morning “Do you know how to knit?” I said immediately, very surprised, “Yes.” Then he said, “Leave all your work aside and knit for me an electrical mantle for the bell jar on the Sequenator so that we can raise the temperature in the cup to 50°C and be able to perform the sequence.” I left everything, I went into the city and bought a pair of gloves, white glass wool and a pair of needles. It was quite strange to sit in the lab and knit the bell jar pattern with a little hole in one sideto enable observation of the reaction inside. The electric wire was inserted between the two knitted layers and it was used in the Sequenator until it was replaced by chemicothermo reaction by spraying the bell jar with stannic chloride.

On another occasion, I remember the difficulty that I encountered when developing and doing thin layer chromatography. Instead of round PTH spots I was getting a boomerang type shape. I was worried at the time with these results and I went to Pehr Edman asking him how to solve this problem. He laughed and turning to me said, “What are you expecting? We are in Australia and it is legitimate to have the boomerangs; try to solve this problem yourself.” It was a difficult task, because this kind of shape was not always seen on the chromatograms and the reason was hard to guess. But eventually after hard work and perseverance, I found out the Camag Kieselgel F254 was not uniformly ground. Contacting Camag factory in Switzerland, this mechanical problem was solved and from then on it did not happen again.

As a scientist, Pehr Edman was characterised by faithful perseverance in pursuing an established goal through years of work. The development of the Sequenator took approximately four to five years. When it was completed, I remember that one day around the table at our usual team time, he asked us what we thought the name of the new machine which would revolutionise the protein chemistry field should be. Names came out and finally I remember that we decided to choose between three names: Matilda (popular song and proposed Australian anthem), Sequana or Sequenator. After a few days, he told us that he thought that the most adequate name would be Sequenator, to describe the work it performs.

As a human being, he was a very sensitive and quiet man. He was a very lonely person, but endeavoured to hide it. His life was enlightened when he got remarried to the charming Agnes née Henschen.

His long hours of loneliness were filled with classical music, books, walks and fishing on the beautiful coastline at Wilson’s Promontory or Apollo Bay. I remember going sometimes to these places with one of his three good friends, Professor Trikojus, Dr Radick and Dr J. Hohne.

He wrote a chapter “Sequence Determination” in the book Protein Sequence Determination, 1975, in which he described some novel modifications of the establishment of the amino acid sequences of peptides and proteins (see box on Further Reading, in the accompanying article). At his request, all the experimental parts of his chapter were thoroughly repeated in his own laboratory by myself, so that any problem encountered would be solved subject to his standards. Pehr Edman was only to see the fruits of his great achievement in part, but his pupils will receive his heritage and will do the best to further his work in Biochemistry.

Elizabeth Minasian worked with Professor Pehr Edman for eight years at St Vincent’s School of Medical Research and since 1970 with Professor Emeritus Syd J. Leach AO at the Biochemistry Department of Melbourne University.