CMLS, Cell. mol. life sci. 53 (1997) 780–789
1420-682X/97/100780-10 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 1997

CMLS Cellular and Molecular Life Sciences

Synthesis, structure and function of poly-α-amino acids – the simplest of protein models

E. Katchalski-Katzir

Weizmann Institute of Science, Department of Membrane Research and Biophysics, Rehovot 76100 (Israel), Fax +972 8 9468256, e-mail: bfkatzir@weizmann.weizmann.ac.il

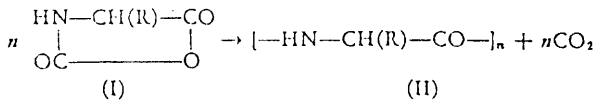
Abstract. During the 1950s, linear and multichain poly-α-amino acids were synthesized by polymerization of the corresponding N-carboxyamino acid anhydrides in solution in the presence of suitable catalysts. The resulting homo- and heteropolymers have since been widely employed as simple protein models. Under appropriate conditions, poly-α-amino acids, in the solid state and in solution, were found to acquire conformations of an α-helix and of β-parallel and antiparallel pleated sheets, or to exist as random coils. Their use in experimental and theoretical investigations of helix-coil transitions helped to shed new light on the mechanisms involved in protein denaturation. Conformational fluctuations of peptides in solution were analysed theoretically and studied experimentally by nonradiative energy-transfer techniques. Poly-α-amino acids played an important role in the deciphering of the genetic code. In addition, analysis of the antigenicity of poly-α-amino acids led to the elucidation of the factors determining the antigenicity of proteins and peptides. The synthetic procedures developed made possible the preparation of immobilized enzymes which were shown to be of considerable use as heterogeneous biocatalysts in the chemical and pharmaceutical industry. Interest in the biological and physicochemical characteristics of poly-α-amino acids was recently renewed because of the reported novel findings that some copolymers of amino acids are effective as drugs in multiple sclerosis, and that glutamine repeats and reiteration of other amino acids occur in inherited neurodegenerative diseases.

Key words. Poly-α-amino acids; α-helix; β-pleated sheets; helix-coil transitions; conformational fluctuations; genetic code; antigenicity; immobilized enzymes; multiple sclerosis; glutamine repeats.

Introduction

During my scientific career, I did not have the privilege of collaborating with Professor Linderstrøm in Copenhagen. I have, however, followed his work with keen interest, and it is because of him and my teachers Professor Max Frankel at the Hebrew University in Jerusalem, Professor John Edsall at the Harvard Medical School in Boston and Professor Herman Mark at the Polytechnic Institute of Brooklyn, that I became enchanted with the study of biopolymers in general and proteins and peptides in particular. Even as a student I realized that proteins, nucleic acids and polysaccharides play a major role in determining life processes, and thus decided to do my best to contribute towards a better understanding of the structure-function relationships of proteins.

While working at the Hebrew University in Jerusalem during the early 1950s it occurred to me that if I could prepare synthetic high-molecular-weight polypeptides consisting of one or a few amino acid residues, a study of their properties might contribute to the elucidation of the structure-function relationships in proteins. While searching the literature for a suitable monomer, I accidentally came across the work of Leuchs in Abderhalden's Handlexikon [1]. After synthesizing N-carboxyglycine anhydride by cyclization of N-carbome-



Scheme 1. Preparation of a linear poly-α-amino acid (II) by the polymerization of an N-carboxy-α-amino acid anhydride (I).

thoxyglycyl chloride, Leuchs had found that it readily gives off carbon dioxide to yield what he called an anhydroglycine. Suspecting that this product was in fact a linear polyamino acid, I decided to synthesize some additional N-carboxy-α-amino acid anhydrides (I), study their polymerization, and characterize the corresponding amino acid polymers obtained (II) (scheme 1). The amine-initiated polymerization of N-carboxy-α-amino acid anhydride is presented in scheme 2.

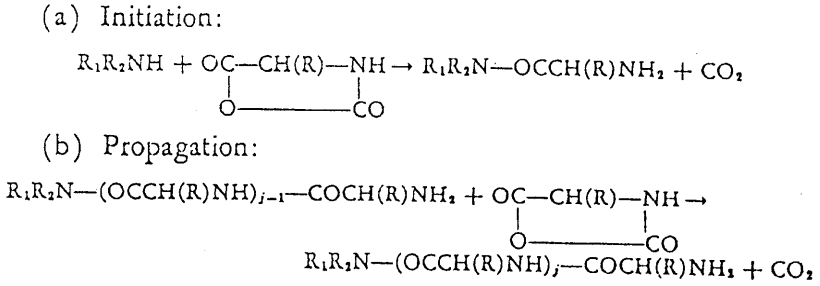
Synthesis of linear and multichain poly-α-amino acids

At first, I was particularly interested in preparing a high-molecular-weight water-soluble polyamino acid, and therefore decided to start with the synthesis of poly-L-lysine (VI). This polypeptide was finally obtained by polymerization of ε,N-carbobenzyloxy-α,N-carboxyL-lysine anhydride (IV) to yield poly-ε,N-carbobenzyloxy-L-lysine (V), followed by removal of the protecting group by suitable means [2] (see scheme 3). Removal of the protecting benzyloxy carbonyl group was extremely difficult and after many attempts I discovered, together with my first Ph.D. student Izhak Grossfeld, that it could be done with PH_4I. At first we assumed that the benzyl groups of the benzyloxycarbonyl residue are reduced by the liberated PH_3; however, since we found ourselves weeping copiously during synthesis we realized that benzyl iodide was being evolved as a result of the HI liberated. These findings led, many years later, to the development of the classical technique for removal of the benzyloxycarbonyl protecting groups with HBr in glacial acetic acid by Arieh Berger and Dov Ben Ishai in my laboratory.

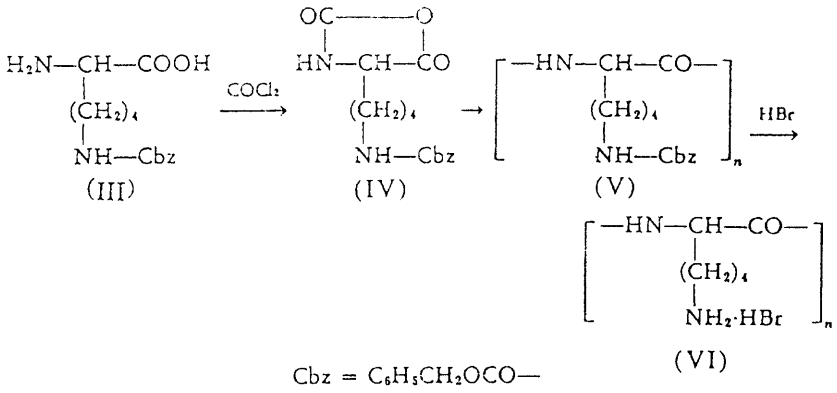
After moving to the Weizmann Institute, I continued to extend my work on polyamino acids as protein models. With my colleagues and students I synthesized a number of other polyamino acids including poly-L-arginine [3], poly-L-histidine [4], poly-L-aspartic acid [5], poly-L-tyrosine [6], poly-L-serine [7], poly-L-cysteine [8], poly-L-proline [9] and poly-L-hydroxyproline [10]. The synthesis of poly-L-glutamate by the polymerization of γ-benzyl-α,N-carboxy-L-glutamate anhydride is illustrated in scheme 4. We also synthesized amino acid copolymers and multichain polyamino acids, i.e. polyamino acids in which a suitable polypeptide backbone, such as polylysine, serves for the attachment of a large number of polypeptide chains, resulting in a branched macromolecule [11] (scheme 5).

Conformation of poly-α-amino acids in the solid state and in solution

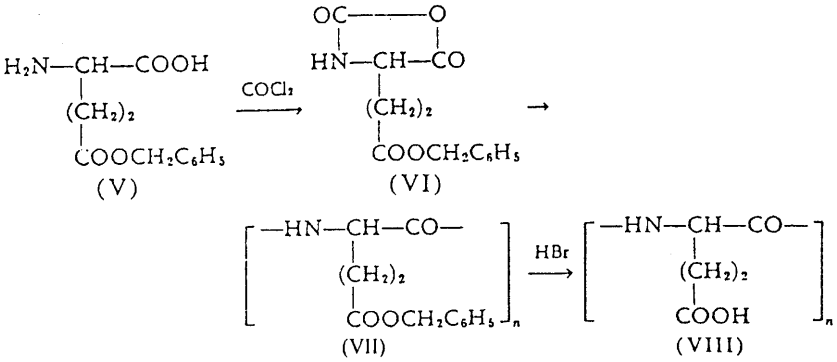
The availability of high-molecular-weight polyamino acids opened the way to a thorough investigation of their conformation, particularly their secondary structure in the solid state and in solution. In 1951, Max Perutz analysed the X-ray diffraction pattern of poly-γ-benzyl-L-glutamate fibres [12] and confirmed the presence of the α-helical polypeptide backbone predicted by Pauling and Corey. Of particular interest was the detection of the 1.50 Å meridianal reflection, corresponding to the residue translation of 1.50 Å along the fibre axis, characteristic of the α-helix. Detailed analysis of the X-ray diffraction pattern of oriented fibres of poly-L-alanine (which acquired an α-helix conformation) was carried out in 1959 by Elliott and Malcolm [13], and their findings were of considerable help to John Kendrew and Max Perutz in their deciphering of the X-ray patterns of myoglobin and haemoglobin. A remarkable observation was that fibres of α-poly-L-alanine undergo a total $\alpha \rightarrow \beta$ transformation when stretched in steam.



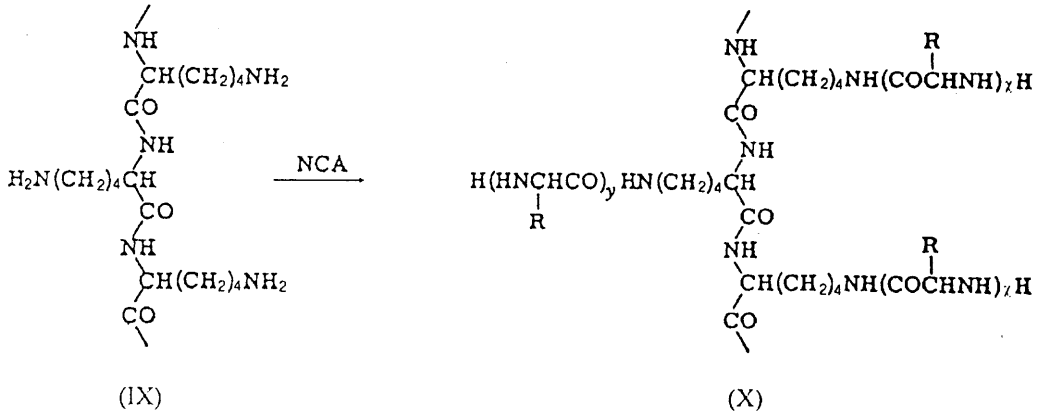
Scheme 2. Amine initiated polymerization of N-carboxy-α-amino acid anhydrides.



Scheme 3. Synthesis of poly-L-lysine (VI) by the polymerization of ε,N-carbobenzyloxy-α,N-carboxy-L-lysine anhydride (IV) and removal of the protecting Cbz-groups with HBr.



Scheme 4. Synthesis of poly-L-glutamic acid (VIII) by the polymerization of γ,benzyl-α,N-carboxy-L-glutamate anhydride (VI), and removal of the benzyl protecting groups of poly-γ,benzyl-L-glutamate (VIII) with HBr.



Scheme 5. Synthesis of multichain poly-amino acids (X) by initiating the polymerization of N-carboxy-α-amino acid anhydrides with poly-L-lysine (IX).

Polyamino acids synthesized and studied in our laboratory by Arieh Berger, Joseph Kurtz and Jurgen Engel, in particular poly-L-proline [9], polyhydroxy-L-proline [10] and the sequential poly(Pro-Gly-Pro) [14], were useful as model compounds in the elucidation of the three-dimensional structure of the fibrous protein collagen. I was especially intrigued by the finding of Kurtz that marked optical mutarotation occurs when our

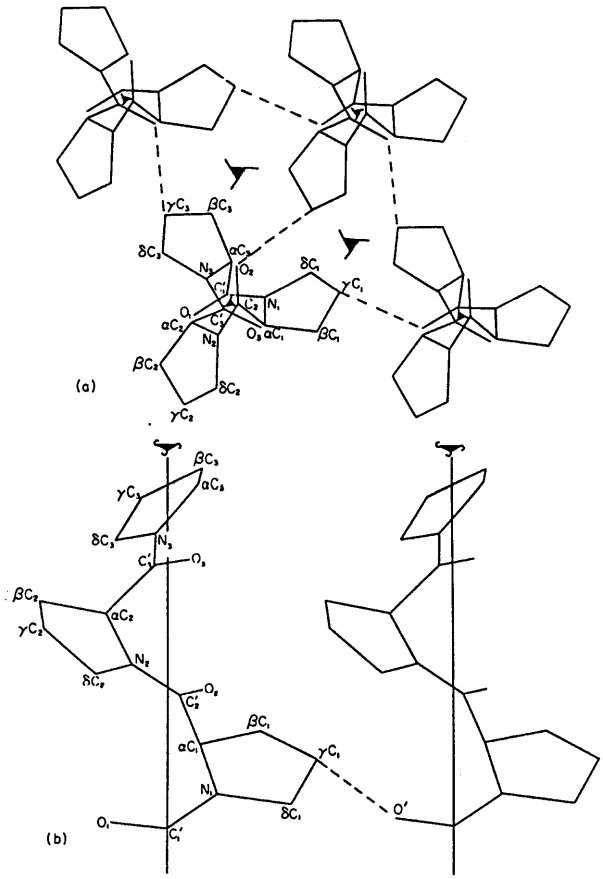


Figure 1. Conformation of poly-L-proline II.

water-insoluble poly-L-proline II, consisting of open right-handed helices with all peptide bonds in the *cis* configuration, is transformed in suitable solvents into the water-soluble poly-L-proline II, consisting of open left-handed helices with all peptide bonds in the *trans*-configuration [15] (fig. 1). In 1954 Ramachandran and Kartha, and later Alex Rich and Francis Crick, suggested that collagen fibres have a triple-stranded helical conformation. Poly(Pro-Gly-Pro) was the first polymer that both in solution and in the solid phase was shown to form the triple-stranded helical conformation attributed to collagen, and was therefore used by chemists and biologists as a model compound for collagen. A comparison between the X-ray diffraction photographs of fibres of $(Gly\text{-}Pro\text{-}Hypro)_n$ and of collagen is given in figure 2.

A considerable amount of work, both experimental and theoretical, was done on the conformation and conformational transitions of polyamino acids not only in the solid state but also in solution. Studies with solvent systems were initiated experimentally by P. Doty and E. Blout and their collaborators at Harvard, and theoretically by W. Moffitt and J. G. Kirkwood and by J. A. Schellman [16, 17]. The existence, in appropriate solvent systems, of regular macromolecular conformations, as well as the existence of the random-coil conformation, was established experimentally by Ignacio Tinoco, J. T. Young, Gerald Fasman and Bruno Zimm by hydrodynamic, optical, electrical, and nuclear magnetic resonance methods [16, 17]. Moreover, conditions were established for attaining a conformational transition between the above macromolecular structures. Taken together, these findings led to a deeper insight into protein denaturation and renaturation. Of particular importance was the finding that the optical rotatory dispersion of poly-α-amino acids in the far ultraviolet (180–260 mμ) in helical conformation differs markedly from that in the random coil form. The helical form is characterized by a large positive Cotton effect and the

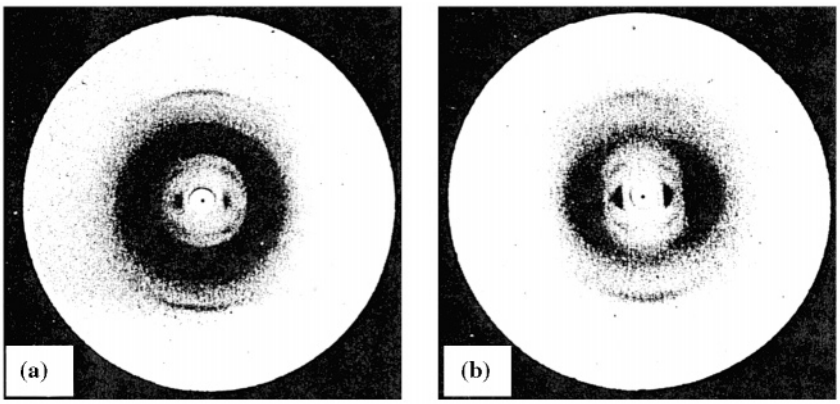


Figure 2. X-ray diffraction photographs of (*a*) fibres of $(Gly\text{-}Pro\text{-}Hypro)_n$ and (*b*) collagen fibres (rat-tail).

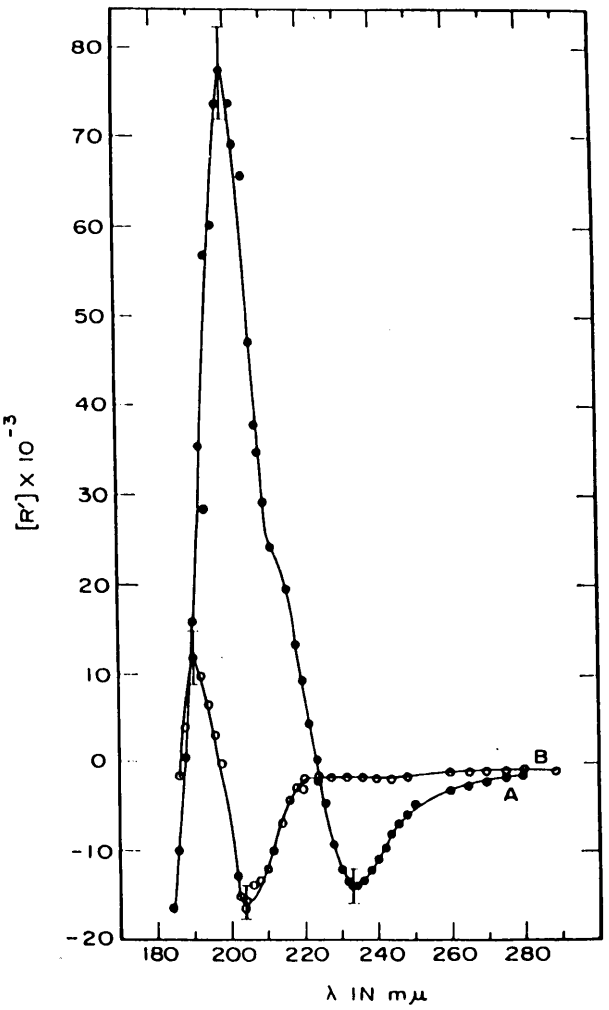


Figure 3. Ultraviolet rotatory dispersion of the helical and random coil forms of poly-α,L-glutamic acid. Curve A (●-●-●): helical form in water solution pH 4.3. Curve B (○-○-○): random coil form (sodium salt), pH 7.1 in water solution [19].

random coil by a smaller negative Cotton effect. The characteristic optical rotatory dispersion of poly-α-L-glutamic acid in its two conformations is given in figure 3 [18]. Careful analysis of the above data permits the prediction of the percentage of amino acid residues in an α-helix or in random coil conformation from the corresponding rotatory dispersion data.

Determination of distance distributions and conformational fluctuations by nonradiative energy-transfer techniques

Because many of the characteristics of linear polymer solutions are determined by their chain length and flexibility, I thought that partially ordered or flexible structures in peptides and globular proteins could also be described in terms of intramolecular distance distributions and rates of transitions and that these could be determined by a time-resolved long-range dynamic nonradiative energy-transfer technique based on Förster's theory [19]. First, with Izchak Steinberg, I analysed the role of diffusion in nonradiative energy-transfer as well as in fluorescence quenching and chemical reactions [20]. Together with Elisha Haas, we then searched for model oligopeptides in which we could use the energy-transfer technique to determine the distribution of end-to-end distances as well as the Brownian motion of the ends of oligopeptide chains in solution. We synthesized a homologous series of oligopeptides, each consisting of four to nine N^5-(2-hydroxyethyl)-L-glutamine residues and containing at its ends a fluorescent donor and a fluorescent acceptor of electronic excitation energy. The chromophores naphthalene and dansyl, used as donor and acceptor respectively, fulfilled the conditions necessary for energy transfer according to the Förster mechanism. The kinetics of fluorescence decay of the donor in a highly viscous glycerol solution enabled the characteristic end-to-end distribution function between the donor and acceptor to be derived [21]. Subsequent analysis of the fluorescence decay curves of these oligopeptides in solvents of low viscosity enabled the apparent diffusion rate of the molecular ends relative to one another to be estimated [22, 23].

The nonradiative energy-transfer technique was subsequently extended by Elisha Haas and his students to different proteins specifically labeled with a fluorescent donor and fluorescent acceptor, enabling them to determine intramolecular distances as well as intramolecular dynamics in various proteins such as pancreatic trypsin inhibitor [24] and phosphoglycerate kinase [25]. Further theoretical and experimental developments have yielded new information on the denaturation and renaturation of proteins and can be expected to elucidate the mechanisms by which proteins fold during their biosynthesis on ribosomes, as well as the preferential folding of peptide hormones upon binding to their corresponding receptors.

Biological properties of poly-α-amino acids

Meanwhile, in Rehovot, I concentrated on the study of the biological properties of polyamino acids [26]. To my delight poly-L-lysine, as well as other homopolyamino acids and amino acid copolymers, turned out to be excellent models for investigating the mechanism of enzymatic protein hydrolysis and transpeptidation. I still remember the excitement with which I followed the rapid hydrolysis of poly-L-lysine by trypsin, using the cumbersome old Van Slyke apparatus. We showed further that the specificity of an enzyme acting on a

Table 1. Guinea pigs with induced experimental allergic encephalomyelitis (EAE) treated with the basic amino acid copolymer, Cop I, consisting of L-alanine, L-glutamic acid, L-lysine and L-tyrosine (at residue molar ratio of 6.0, 1.9, 4.7, 1.0, respectively).

	Guinea pigs with initial attack of EAE	Average day of onset of disease	Guinea pigs with relapses of EAE	Severity of relapses
Control group	10/10 (5 died)	11.4 (severe)	5/5	very high
Suppression by Cop I	8/8	13.2	6/8	moderately mild
Prevention Cop I	5/12	89.8	1/12	very mild

high-molecular-weight polypeptide is often strikingly different from that observed with low-molecular-weight peptides. Partial hydrolysis of poly-L-lysine yields, as expected, a mixture of lysine oligomers. These were separated chromatographically and investigated immunologically by my former student, Arieh Yaron, in Herb Sober's laboratory at NIH. Penetration of these oligomers into *Escherichia coli* was studied by Charles Gilvarg [27], a visiting scientist at the Weizmann Institute. By using a lysineless mutant of *E. coli* Gilvarg was able to show that all oligomers up to tetralysine, but no higher, can penetrate readily into *E. coli* and permit growth of the lysine auxotroph in their presence.

In our experiments with a prolineless mutant of *E. coli*, Sara Sarid observed that the organism can grow on a synthetic medium in which poly-L-proline is substituted for L-proline. Clearly, the polymer was being hydrolysed by an unknown enzyme. Further studies by Arieh Yaron on the cleavage of various synthetic proline-containing oligo- and polypeptides led to the identification and characterization of the enzyme aminopeptidase P present in pro- and eukaryotes [28].

Poly-α-amino acid antigenicity

An important outgrowth of the studies on synthetic polyamino acids was the development in my laboratory of techniques for the preparation of polypeptidyl proteins, i.e. proteins to which polypeptide chains are covalently attached via amide bonds to the free amino groups of the protein. The synthesis of polytyrosyl gelatin and the demonstration that it is antigenic, whereas the unmodified protein is not [29], led in 1960 to the preparation by Michael Sela and Ruth Arnon, then in my department, of the first fully synthetic antigen. In this compound, tyrosine and glutamic acid residues are attached to a multipoly-DL-alanyl poly-L-lysine [30, 30a]. I vividly remember our immunological experiments, in which guinea pigs injected two or three times with polytyrosyl gelatin went into anaphylactic shock – a most unpleasant experience for the guinea pigs and a sobering demonstration to me of how careful one should be in treating living beings with synthetic or even native polymers. The way was opened for the fundamental and extensive studies of Sela and his coworkers on the chemical and genetic basis of antigenicity [31]. Some of the polypeptidyl enzymes we prepared retained full enzymatic activity. This finding was the basis for our subsequent preparation of a great variety of immobilized enzymes, which are of theoretical and practical interest [31–34].

Immobilized enzymes

My interest in enzyme-polymer conjugates was aroused by the growing body of data indicating that many of the enzymes embedded in organelles or biological membranes within cells act as heterogeneous catalysts. An interesting experiment was to immobilize artificially enzymes and study their properties, especially their kinetic characteristics, under controlled conditions. Moreover, I felt that such immobilized enzymes could be utilized in the construction of novel enzyme reactors of use in the laboratory, the clinic and industry.

My first paper on a water-insoluble enzyme was published in 1960 [32], in which I described the preparation of a water-insoluble trypsin derivative and its use in a trypsin column. The method involved preparation of a polytyrosyl trypsin derivative containing tyrosyl peptide side chains by means of the polymerization of *N*-carboxy-L-tyrosine anhydride with trypsin [35] and coupling of the resulting water-soluble derivative with a diazotized copolymer of ρ-amino phenylalanine and leucine to yield the required water-insoluble trypsin. The trypsin column showed high activity towards benzoyl-L-arginine methyl ester, poly-L-lysine, and other well-known synthetic and native trypsin substrates. Of particular interest was the finding that the enzymatic activity of the water-insoluble trypsin remained practically unaltered in dilute HCl at 2 °C. Immobilization prevented autodigestion, and blocking of the ε-amino groups of the enzyme led to a marked decrease in the number of peptide bonds susceptible to trypsin.

These encouraging results prompted us to prepare other immobilized enzymes [33] such as immobilized chymotrypsin, urease, papain, alkaline phosphatase, and carboxypeptidase, in each case by covalent binding of the enzyme via nonessential side groups to water-insoluble carriers.

Table 2. Uninterrupted repeats of ≥10 codons.

Reiterant AA	Codon	Access No.	Gene	Species
Ala	$(GCA)_{11}$	M98269	antho-RFamide neuropeptide	*A. elegantissima*
Asn	$(AAC)_{20}$	X16523	AAC-rich mRNA (pl.K330)	*D. discoideum*
	$(AAT)_{16}$	S55235	cAR3 = cAMP receptor subtype 3	*D. discoideum*
	$(AAT)_{14}$	M87278	adenyl cyclase germination protein	*D. discoideum*
Asp	$(GAT)_{14}$	M60052	histidine-rich calcium binding protein	*Homo sapiens*
Gln	$(CAA)_{11}$	L19349	hydroxymethylglutaryl CoA reductase	*D. discoideum*
	$(CAA)_{20}$	M17826	SSN6 or CYC8	*S. cerevisiae*
	$(CAA)_{15}$	M60807	merozoite surface antigen I	*P. falciparum*
	$(CAG)_{21}$	L12392	Huntington's disease gene	*Homo sapiens*
	$(CAG)_{19}$	Y00489	ventral prostate glucocorticoid receptor	*Rattus rattus*
	$(CAG)_{19}$	M55654	TATA-box binding protein (TBP)	*Homo sapiens*
Glu	$(GAA)_{10}$	J03998	glutamic acid-rich protein	*P. falciparum*
	$(GAC)_{12}$	J05080	histidine-rich calcium-binding protein	*O. cuniculus*
Gly	$(GGA)_{10}$	M18289	E1B large T-antigen	*Adenovirus 41*
	$(GGC)_{20}$	M23263	androgen receptor	*Homo sapiens*
His	$(CAT)_{11}$	S55234	cAMP receptor subtype 2	*D. discoideum*
Ser	$(AGC)_{10}$	X15898	sporozite antigen	*Eimeria tenella*
	$(AGC)_{10}$	M88749	vitellogenin	*I. unicuspus*
	$(AGT)_{11}$	M31431	attachment protein	*Myco. genitalium*
Thr	$(ACA)_{17}$	M66619	aminocyclopropane carboxylate synthesis	*D. caryophyllus*

Growing interest in immobilized enzymes led to the development by various groups of novel enzyme immobilization techniques in which enzymes were adsorbed or covalently bound to organic or inorganic carriers, or entrapped in gels, fibres or microcapsules, and systems in which enzymes remained in solution but functioned in a limited space enclosed by an ultrafiltration membrane [36]. In a novel enzyme immobilization technique that I developed together with my collaborators at Tel Aviv University's Biotechnology Center, immobilized monoclonal antibodies were used as carriers to combine with their corresponding enzyme antigen [37]. With this technique, immobilization did not result in any loss of enzymatic activity.

Thus within a relatively short period, we obtained a great variety of immobilized enzymes as well as enzyme reactors of various types [38] which opened the way for the use of immobilized enzymes in the food, pharmaceutical and chemical industries.

The first industrial use of immobilized enzymes was reported in 1967 by Chibata and coworkers of the Tanabe Seiyaku Company in Japan, who developed columns of immobilized *Aspergillus oryzae* aminoacylase for the resolution of synthetic racemic DL-amino acids into the corresponding optically active enantiomers. Around 1970, two other immobilized systems were launched on a pilot plant scale. In England, immobilized penicillin acylase, also referred to as penicillin amidase, was used to prepare 6-amino-penicillanic acid from penicillin G or V; and in the USA, immobilized glucose isomerase was used to convert glucose into fructose. These successful industrial applications prompted extensive research in enzyme technology, leading to a steady increase in the number of industrial processes based on sophisticated immobilized enzyme reactors.

The use of immobilized enzymes in industry is now well established [33]. The intermediate compound 6-amino-penicillanic acid, which is employed in the preparation of the semisynthetic penicillin derivatives used as oral antibiotics, is now prepared worldwide by means of an immobilized enzyme process; its estimated production in 1992 was 7500 tons. In the same year, 15,000 tons of acrylamide were produced from acetonitrile by use of immobilized nitrile hydratase, and the production of high-fructose corn syrup from glucose by immobilized glucose isomerase reached 8 million tons [34].

Use of poly-α-amino acids in the deciphering of the genetic code

Knowledge of the physical and chemical properties of synthetic polypeptides played a decisive role in the work that led in 1961 to the cracking of the genetic code. In their first paper on the subject, Marshall Nirenberg and J. H. Matthei identified the poly-L-phenylalanine produced enzymatically, in a cell-free system in the presence of polyuridylate used as messenger [39], with the poly-L-phenylalanine we had synthesized in Rehovot in 1955. As it happens, Michael Sela was at NIH when Nirenberg was working on the code, and he had been able to inform Nirenberg that the normally insoluble poly-L-phenylalanine could be dissolved in acetic acid saturated with HBr. Soon afterwards Nirenberg identified other homo- and heteropolyamino acids in the deciphering of the genetic code: poly-A was found to code for poly-L-lysine, poly-C for poly-L-proline and poly-G for polyglycine [40].

Cop I, a copolymer of Ala, Glu, Lys and Tyr, as a potential drug against multiple sclerosis

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system, in which infiltrating lymphocytes, predominantly T-cells and macrophages, cause damage of the myelin sheath. It is thought to be an autoimmune disease, probably associated with an early viral infection. Experimental allergic encephalomyelitis (EAE) serves as the experimental animal model for the autoimmune process in MS. In view of the immunological nature of EAE, attempts have been made to suppress the disease in animals challenged with myelin basic protein (MBP). In guinea pigs, for example, high doses of MBP in incomplete Freund's adjuvant were shown to be highly effective in preventing EAE when administered before sensitization, and in suppressing the disease when given after sensitization. These findings led M. Sela and R. Arnon to test the effect on EAE of a basic amino acid copolymer, consisting of L-alanine, L-glutamic acid, L-lysine and L-tyrosine (at residue molar ratios of 6.0, 1.9, 4.7, 1.0, respectively), prepared in my laboratory by the polymerization of the corresponding N-carboxy-amino acid anhydrides. As expected, the copolymer imitated the immunogenic effects of MBP and could suppress the model disease in guinea pigs when injected at suitable concentrations and time intervals [41] (table 1). None of the copolymers tested was found to be toxic or to exhibit any general immunosuppressive activity. These findings prompted Sela and Arnon to test the effect of Cop I on sick monkeys showing signs of EAE, and the encouraging results obtained opened the way for testing the effect of Cop I on MS in humans. Controlled experiments carried out so far have indicated that Cop I represents a potential low-risk drug for the treatment of MS. Analysis of the immunochemical mechanism of EAE suppression by Cop I revealed that Cop I and Cop I-derived peptides can bind to the relevant major histocompatibility complex (MHC) molecules and competitively inhibit the binding of MBP. The microheterogeneity of Cop I might partially account for its success, as it may contain numerous amino acid sequences that can successfully complete with MBP for class II MHC antigens of many different genetic backgrounds.

As a result of the efforts of Sela and Arnon, and their close collaboration with Teva, the largest Israeli pharmaceutical company, it was possible to carry out clinical trials with Cop I, all of which have shown that the copolymer is of help in alleviating the suffering of multiple sclerosis patients. The Food and Drug Administration recently approved its use in the USA. The copolymer appears now in the market under the name Copaxone [42].

Proteins with glutamine repeats and reiteration of other amino acids

Four inherited neurodegenerative diseases are linked to abnormally expanded repeats of glutamine residues near the N-termini of the affected proteins: Huntington disease (HD); spinal and bulbar muscular atrophy (SBMA), also known as Kennedy's disease; spinocerebral ataxia type I (SCAI); and dentatorubral-pallidoluysian atrophy (DRPLA) [43]. All four diseases begin earlier and become more severe as the glutamine repeats become longer. The repeats tend to lengthen in successive generations of affected individuals, especially in male transmission. These findings prompted M. Perutz and his collaborators to construct molecular models of poly-L-glutamine and study their optical, electron and X-ray diffraction properties. Their data revealed the presence of β-sheets strongly held together by hydrogen bonds, suggesting that glutamine repeats may function as polar zippers by joining specific transcription factors bound to separate DNA-segments. Lengthening of the repeats may intensify the disease either as a result of increased nonspecific affinity between such factors or by gradual precipitation of the affected proteins in neurons.

Perutz extended the above studies to other polar zippers in proteins with reiterated sequences of polar amino acids, and attempted to predict their role in human disease [44]. *Ascaris* haemoglobin, for example, consists of eight subunits, each containing a C-terminal peptide with the sequence Glu-Glu-Lys-His repeated four times. When plotted on a β-strand, this sequence leads to alternate lysines and glutamates on one side of the strand and alternate glutamates and histidines on the other side, suggestive of a polar zipper that links the subunits together. A computer search of the protein database showed that the same or similar sequences also occur in other proteins. Some contain long repeats of Asp-Arg or Glu-Arg, among them the small nuclear ribonucleo-U1 70-kD protein, which is an autoantigen in systemic lupus erythematosus. These repeats appear to constitute the dominant epitopes in the autoimmune reaction.

An impressive set of data on codon reiteration was recently publised by H. Green and N. Wang [45] (table 2). In line with the findings discussed above, these data show that hydrophobic amino acids, and particularly glutamine, account for a large proportion of the longer repeats. In the genes for these proteins the most common repeats are those that contain poly (CAG), even out-of-frame or, to a lesser degree, those that contain repeated doublets of CA, AG or GC. A particularly intriguing finding is that the amino acid sequence of the mastermind gene (*mam*) in *Drosophila virilis* has a region (residue 106–161) containing 44 glutamine residues of which 31 are encoded by CAG and 13 by

CAA. All of the remaining 12 amino acid residues are histidines (encoded by CAC or CAT). These interrupt the glutamine codons at ten locations and are probably the result of single nucleotide substitutions in either CAG or CAA codons. Reiterated hydrophobic or basic residues hardly appear in proteins containing 20 or more residues, suggesting that they are poorly tolerated in proteins.

The mechanism of reiteration is still unclear. It should be mentioned, however, that nucleotide repeats have also been observed in regions of DNA which do not code for proteins; their role in the regulation of cellular behaviour has yet to be elucidated.

It is worth noting that in vertebrates there are specialized telomeric structures, located at the ends of eukaryotic chromosomes, that appear to function in chromosome protection, positioning and replication. These telomeres consist of hundreds or thousands of tandem repeats of the sequence TTAGGG [46]. In all normal somatic cells examined to date, terminal restriction fragment analysis has shown that with each cell division the chromosomes lose about 50 to 200 nucleotides of telomeric sequence. This finding led to the suggestion that the shortening of telomeres functions as a mitotic clock by which cells count and ultimately limit their division. Remarkably, all immortal cells examined to date show no net loss of telomeric length with cell division, suggesting that maintenance of telomeres is required in order for cells to escape from replicative senescence.

Concluding remarks

It is now more than 40 years since linear and multichain poly-α-amino acids were first synthesized by polymerization of the corresponding N-carboxy-α-amino acid anhydrides in the solid state or in solution. The resulting linear homo- and heteropolymers were employed as simple protein models to verify the presence of the main secondary conformations of proteins, namely the α-helix and the parallel and antiparallel β-sheets. They were also used in experimental and theoretical investigations of helix-coil and helix-β strand transitions in the solid state and in solution, thus helping to shed new light on the mechanisms involved in protein denaturation.

In addition, poly-α-amino acids were used to elucidate the mode of action of proteolytic enzymes on synthetic macromolecular substrates and to detect new proteolytic enzymes. Some of them played an important role in the deciphering of the genetic code. Analysis of the antigenicity of poly-α-amino acids led to the elucidation of the factors determining the antigenicity of proteins and peptides.

There has been renewed interest in the biological and physicochemical characteristics of poly-α-amino acids because of the recent finding that some copolymers of amino acids are effective as drugs in multiple sclerosis, and that glutamine repeats and reiteration of other amino acids occur in inherited neurodegenerative diseases. The presence of repeating sequences of amino acid in proteins and of nucleotides in DNA raises many interesting questions about their respective roles in determining protein structure and function, and gene performance and regulation. In seeking answers, the poly-α-amino acids, as the simplest protein models, can once again be expected to play a useful role.

1 Abderhalden E. (1911) Biochemisches Handlexikon, vol. 4, 421 pp., Springer Verlag, Berlin
2 Katchalski E., Grossfeld I. and Frankel M. (1947) Poly-lysine. J. Am. Chem. Soc. **69:** 2564–2565
3 Katchalski E. and Spitnik P. (1949) Poly-arginine. Nature **164:** 1092–1097
4 Patchornik A., Berger A. and Katchalski E. (1957) Poly-L-histidine. J. Am. Chem. Soc. **79:** 5227–5230
5 Berger A. and Katchalski E. (1951) Poly-L-aspartic acid. J. Am. Chem. Soc. **73:** 4084–4088
6 Katchalski E. and Sela M. (1953) The synthesis and spectrophotometric study of poly-L-tyrosine and poly-3.5-diodotyrosine. J. Am. Chem. Soc. **75:** 5284–5289
7 Bohak Z. and Katchalski E. (1963) Synthesis, characterization and racemization of poly-L-serine. Biochemistry **2:** 228–237
8 Sadeh T. and Katchalski E. (1960) Poly-DL-homocysteine, preparation and properties. Bull. Res. Counc. Israel **9A:** 65
9 Berger A., Kurtz J. and Katchalski E. (1954) Poly-L-proline. J. Am. Chem. Soc. **76:** 5552–5554
10 Kurtz J., Fasman G. D., Berger A. and Katchalski E. (1958) Poly-hydroxy-L-proline. J. Am. Chem. Soc. **80:** 393–397
11 Sela M. and Katchalski E. (1955) Branched polyamino acids. Experientia **11:** 62–66
12 Perutz M. F. (1951) New X-ray evidence on the configuration of polypeptide chains. Nature **167:** 1053–1054
13 Elliot A. and Malcolm B. R. (1959) Chain arrangement and sense of the α-helix in poly-L-alanine fibers. Proc. Roy. Soc. (London) **A249:** 30–41
14 Engel J., Kurtz J., Katchalski E. and Berger A. (1966) Polymers of tripeptides as collagen models. II. Conformational changes of poly(L-prolyl-glycyl-L-prolyl) in solution. J. Molec. Biol. **17:** 255–272
15 Steinberg I. Z., Sela M., Harrington W. F., Berger A. and Katchalski E. (1960) The configurational changes of poly-L-proline in solution. J. Am. Chem. Soc. **82:** 5263–5279
16 Various authors. (1967) Summarizing articles. In: Poly-α-amino acids – Protein Models for Conformational Studies, Fasman G. D. (ed.), Marcel Dekker Inc., New York
17 Katchalski E., Sela M., Silman H. I. and Berger A. (1964) Polyamino acids as protein models. In: The Proteins, vol. II, 2nd ed., Neurath M. (ed.), Academic Press, New York
18 Blout E. R., Schmier I. and Simmons N. S. (1962) New cotton effects in polypeptides and proteins. J. Am. Chem. Soc. **84:** 3193–3194
19 Forster T. (1948) Zwischen Molekularer Energiewanderung und Fluoreszenz. Ann. Phys. Leipzig **2:** 55–75
20 Steinberg I. Z. and Katchalski E. (1968) Theoretical analysis of the role of diffusion in chemical reactions, fluorescence quenching and nonradiative energy transfer. J. Chem. Phys. **48:** 2404–2410
21 Haas E., Wilchek M., Katchalski-Katzir E. and Steinberg I. Z. (1975) Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer. Proc. Natl Acad. Sci. USA **72:** 1807–1811

22 Haas E., Katchalski-Katzir E. and Steinberg I. Z. (1978) Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends. Biopolymers **17:** 11–31
23 Katchalski-Katzir E., Haas E. and Steinberg I. Z. (1981) Study of conformation and intramolecular motility of polypeptides in solution by a novel fluorescence method. Am. NY Acad. Sci. **366:** 44–61
24 Gottfried D. S. and Haas E. (1992) Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. Biochemistry **31:** 12353–12362
25 Haran G., Haas E., Szpikowska B. K. and Mas M. T. (1992) Domain motions in phosphoglycerate kinase: determination of interdomain distance distributions by site-specific labeling and time-resolved fluorescence energy transfer. Proc. Natl Acad. Sci. USA **89:** 11764–11768
26 Sela M. and Katchalski E. (1959) Biological properties of poly-α-amino acids. Adv. Protein Chem. **14:** 391–477
27 Gilvarg C. and Katchalski E. (1965) Peptide utilization in *E. coli*. J. Biol. Chem. **240:** 3093–3098
28 Yaron A. and Naider F. (1993) Proline-dependent structural and biological properties of peptides and proteins. Crit. Rev. Biochem. Molec. Biol. **28(1):** 31–81
29 Sela M., Katchalski E. and Olitzki A. L. (1956) Anaphylactic shock in guinea pigs sensitized to polytyrosylgelatin. Science **123:** 1129
30 Sela M. and Arnon R. (1960) Specific synthetic polypeptide antigens. Biochim. Biophys. Acta **40:** 382–384
30a Sela M. and Fuchs S. (1964) Antigenicity of some new synthetic polypeptides and polypeptidyl gelatins. Biochem. J. **93:** 566–572
31 Sela M. (1969) Antigenicity: Some molecular aspects. Science **166:** 1365–1374
32 Bar-Eli A. and Katchalski E. (1960) A water-insoluble trypsin derivative and its use as a trypsin column. Nature **188:** 856–857
33 Tanaka A., Tosa T. and Kobayashi T. (eds) (1993) Industrial Application of Immobilized Enzymes, Marcel Dekker, New York
34 Katchalski-Katzir E. (1993) Immobilized enzymes – learning from past successes and failures. Trends in Biotechnology **11:** 471–478
35 Glazer A. N., Bar-Eli A. and Katchalski E. (1962) Preparation and characterization of polytyrosyl trypsin. J. Biol. Chem. **237:** 1832–1838
36 Chibata I. (ed.) (1978) Immobilized Enzymes, Wiley & Sons, New York
37 Solomon B., Koppel R., Pines G. and Katchalski-Katzir E. (1986) Enzyme immobilization via monoclonal antibodies. I. Preparation of a highly active immobilized carboxypeptidase A. Biotech. Bioeng. **28:** 1213–1221
38 Mosbach K. (ed.) (1976, 1987, 1988) Methods in Enzymology, vols. 44, 135B, 136C, 137D, Academic Press, New York
39 Nirenberg M. (1973) The genetic code. In: Nobel Lectures: Physiology or Medicine (1963–1970), pp. 372–395, Elsevier, New York
40 Khorana H. G. (1973) Nucleic acid synthesis in the study of the genetic code. In: Nobel Lectures: Physiology or Medicine (1963–1970), pp. 341–369, Elsevier, New York
41 Sela M. (1993) Polymeric drugs as immunomodulatory vaccines against multiple sclerosis. Makromol. Chem. Macromol. Symp. **70/71:** 147–155
42 Arnon R. (1996) The development of Cop I (Copaxone), an innovative drug for the treatment of multiple sclerosis: personal reflections. Immunol. Lett. **50:** 1–15
43 Perutz M. F., Johnson T., Suzuki M. and Finch J. T. (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc. Natl Acad. Sci. USA **91:** 5355–5358
44 Perutz M. (1994) Polar zippers: their role in human disease. Prot. Sci. **3:** 1629–1637
45 Green H. and Wang N. (1994) Codon reiteration and the evolution of proteins. Proc. Natl Acad. Sci. USA **91:** 4288–4302
46 Kim N. W., Piatyszek M. A., Prowse K. R., Harley C. B., West M. O., Ho P. L. C. et al. (1994) Specific association of human telomerase activity with immortal cells and cancer. Science **266:** 2011–2015