

Role of proteolytic enzymes in biological regulation (A Review)*

(limited proteolysis/zymogen activation/control mechanisms)

HANS NEURATH AND KENNETH A. WALSH

Department of Biochemistry, University of Washington, Seattle, Wash. 98195

Contributed by Hans Neurath, September 7, 1976

ABSTRACT Many enzymes, hormones, and other physiologically active proteins are synthesized as inactive precursors (zymogens) that are subsequently converted to the active form by the selective enzymatic cleavage (limited proteolysis) of peptide bonds. The ultimate agency of activating enzymatic function is limited proteolysis, either in a single activation step or in a consecutive series (cascade). The specificity of each activation reaction is determined by the complementarity of the zymogen substrate and the active site of the attacking protease. The sequence of consecutive activation reactions is regulated by the specificity of each enzyme, whereas the degree of amplification of the initial stimulus is determined by the efficiency of each activating step.

Zymogen activation produces a prompt and irreversible response to a physiological stimulus, and is capable of initiating new physiological functions. Typical examples are the processes of blood coagulation, fibrinolysis, complement activation, hormone production, metamorphosis, fertilization, supramolecular assembly, and digestion. The zymogens of the pancreatic serine proteases, in particular, have served as models for detailed studies of the nature of the molecular changes that are involved in the dramatic increase in enzymatic activity that ensues upon limited proteolysis of the zymogen.

In recent years, it has become evident that many proteins are synthesized as inactive precursors or zymogens and that these are subsequently converted to physiologically active forms by the selective enzymatic cleavage of peptide bonds. This process is known as zymogen activation, a term which initially was applied to the activation of precursors of proteolytic enzymes such as trypsinogen, chymotrypsinogen, or procarboxypeptidase (1). It is now apparent that the same type of reaction is involved in a great variety of biological processes, such as blood coagulation, fibrinolysis, complement reaction, hormone production, development, differentiation, and supramolecular assembly, all of which involve zymogen activation in one or more steps (2-9). In the present article, we shall attempt to show that activation by limited proteolysis is indeed an important control element which can initiate new physiological functions or regulate preexistent ones.

Virtually all zymogen activation reactions require the enzyme-catalyzed cleavage of a unique peptide bond by "limited proteolysis." This term was first introduced by Linderstrøm-Lang and Ottesen (10) to describe the restrictive peptide bond cleavage that induces the conversion of ovalbumin to a different crystalline form, plakalbumin, under the influence of the bacterial protease subtilisin. Numerous examples of limited proteolysis have since been described and studied in detail, such as the tryptic conversion of chymotrypsinogen to chymotrypsin (1), the release by subtilisin of the amino-terminal segments of ribonuclease (11), and the conversion of proinsulin to insulin (5). Limited proteolysis is the last step in the synthesis of many

biologically active proteins and probably the first step in protein degradation (12). The specificity of limited proteolysis is best understood in terms of the three-dimensional structure of a protein substrate and of the attacking protease because the region of the protein substrate containing the susceptible peptide bond must fit the active site of the attacking protease in order for amino acid residues of the substrate to interact with primary as well as secondary binding sites of the enzyme (13). In general, limited proteolysis is therefore directed toward surface loops and random segments of polypeptide chains rather than toward internal domains, helices, or pleated sheets.

The activation of zymogens usually occurs by proteolytic cleavage of a peptide bond in a region that is amino terminal relative to the active site of the protein. This may be a consequence of the process of protein biosynthesis, which proceeds in the direction from the amino to the carboxyl end. If it is assumed that the protein assumes its correct tertiary structure as regions of the polypeptide chain are synthesized, the zymogen will be formed prior to the enzyme. Were the activation peptide attached to the carboxyl end, trypsin would be synthesized before trypsinogen, fibrin before fibrinogen, or collagen before procollagen. By synthesizing an inactivating prefix before synthesizing the active portion of the protein molecule, premature physiological function is avoided.

The position of zymogen activation in the overall scheme of physiological control processes is diagrammatically shown in Fig. 1. The term *zymogen* is being used herein to denote in general an inactive precursor that can be converted to an active protein by the cleavage of one or more peptide bonds. This process is essentially irreversible because, in common with many other hydrolytic reactions, proteolysis is an exergonic reaction under normal physiological conditions and there are no simple biological mechanisms to repair a broken peptide bond. In this respect, zymogen activation differs in kind from the freely reversible mechanisms of allosteric transition or covalent modification (14). Whereas the latter are suited to maintain or modulate a steady state of intermediary metabolites, zymogen activation, by virtue of its operational irreversibility, can effect unidirectional changes in the cellular environment and can induce new physiological functions. This type of initiation is more rapid than that regulated by the selective transcription of a genome and is triggered by signals that operate entirely on the post-translational level. Typical zymogen activation reactions are summarized in Table 1.

In some of these processes, the zymogen is converted to the active protein in a single step, whereas in others the process involves consecutive steps or cascades (2) which serve to amplify small stimuli to major physiological responses. Many zymogen activation reactions may have remained undetected thus far because the precursor becomes activated prior to isolation. Indeed, isolation procedures are usually designed for maximum yield of active protein rather than of zymogen and thus may contravene the demonstration of a zymogen precursor.

* By invitation. From time to time, reviews on scientific and technological matters of broad interest are published in the PROCEEDINGS.

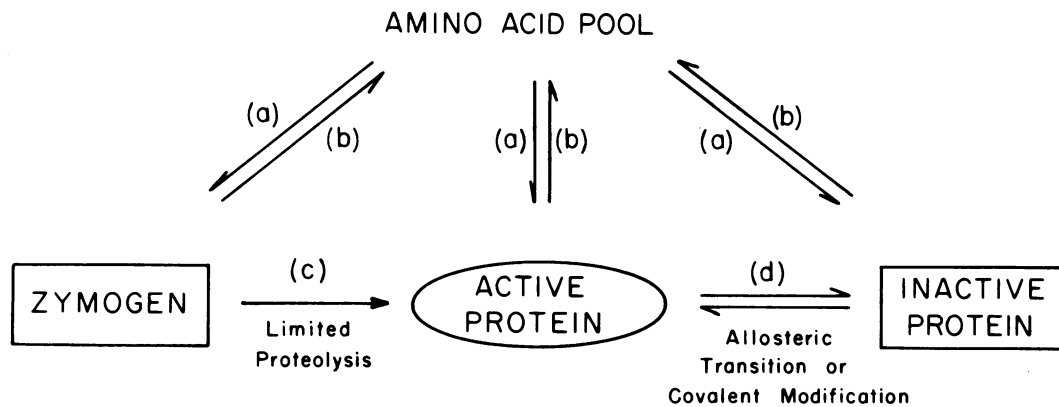


FIG. 1. Schematic representation of major control mechanisms. (a) Transcription and translation regulate the rate of formation of the various proteins from the amino acid pool. (b) Other controls regulate the rate of degradation of the various proteins to their constituent amino acids. (c) The activity of the protease that catalyzes zymogen activation may be in turn regulated by a series of consecutive reactions of limited proteolysis (Fig. 2). Activation of a zymogen is essentially irreversible *in vivo*. (d) Reversible conformational changes are responsive to effector concentrations or to the activities of specific group transferases and hydrolases.

CONSECUTIVE ZYMOGEN ACTIVATION REACTIONS

A series of consecutive zymogen activation reactions is shown diagrammatically in Fig. 2. X, Y, and Z are zymogens, each having the potential of being converted to an active protein. Conversion of the zymogen X to the protease X_a is triggered by a specific physiological stimulus; in the ensuing cascade, the product of one reaction is a catalyst for the next. The sequence of the events is determined by the specificity of each enzyme and the degree of amplification of the initial stimulus is determined by the efficiency of each activating step. For instance, one molecule of X_a might produce 10^3 molecules of Y_a , which in turn produce 10^6 molecules of active protein. When the zymogen is produced by one cell type and the activating protease by another, communication between the two cell types adds another element to the control mechanism (15). For instance, the activation of pancreatic trypsinogen, which originates in the acinar cells of the pancreas, is triggered by enterokinase, which is secreted from the brush border of the small intestine (16). It should be noted, parenthetically, that the activation of

trypsinogen by enterokinase is so specific that only a single bond out of 228 in trypsinogen is cleaved and no other protein has yet been reported to be a substrate for enterokinase (16). Thus, the site of generation of active trypsin is restricted to the confluence of these two secretory streams (Fig. 3). Active trypsin in turn catalyzes the conversion of other pancreatic zymogens to their active forms, i.e., the chymotrypsinogens, proelastase, the procarboxypeptidases, and prophospholipase. This system constitutes a two-stage cascade.

A more complex and extensive cascade system is found in the blood coagulation process, shown in Fig. 4. In fact, the term *cascade* or *waterfall* was introduced by Macfarlane (17) and by Davie and Ratnoff in 1964 (18) in connection with this series of reactions. Five known proteolytic reactions occur along the so-called intrinsic pathway, which is mediated entirely by components found in the plasma, and four proteolytic steps accompany the extrinsic pathway, which includes factors found in tissues (3). The intrinsic and the extrinsic pathways converge to produce thrombin, which in turn converts fibrinogen to fibrin.

Table 1. Examples of zymogens that are converted to active proteins in response to various stimuli

Normal physiological response	Response to foreign stimulus	Programmed response in development or repair
Vasoactive products	Digestion	Development
Angiotensinogen (61)	Pepsinogen	Prochitin synthetase (7)
Prekallikrein*	Prochymosin	Prococoonase (30)
Kininogen	Trypsinogen*	Procollagenase (68)
Hormones	Chymotrypsinogen	Proacrosin (31)
Proinsulin	Procarboxypeptidases	Fibrinolysis
Proglucagon (62)	Prophospholipase (67)	Plasminogen proactivator*
Parathyroid hormone	Proelastase	Plasminogen
Enzymes	Blood coagulation	Toxin
Protyrosinase (63, 64)	Factors VII*, IX*, X*	Promellitin (69)
Prephenoloxidase (65)	Prothrombin*	Self-assembly
Prorenin (66)	Factor XIII	Procollagen
	Fibrinogen	Phage head proteins
	Complement	Processing of zymogen precursors
	Properdin precursor	(34-37)
	Factors C ₃ , C ₅ *	

Except where indicated by a specific reference, references to these zymogens are found in any of three general sources: *Proteolytic Enzymes*, Methods in Enzymology, Vol. 19 (77); *Peptide Bond Hydrolysis*, The Enzymes, Vol. 3 (78); and *Proteases and Biological Control*, Cold Spring Harbor Conferences on Cell Proliferation, Vol. 2 (79).

* An intermediate component in a series of consecutive zymogen activations.

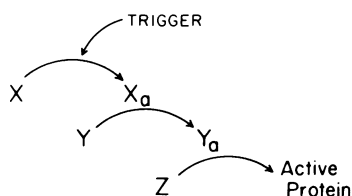


FIG. 2. Schematic representation of consecutive zymogen activation reactions (cascade). Activation of the zymogens X, Y, and Z leads to the formation of the proteases X_a and Y_a and the active protein, in response to a trigger which initiates the conversion of X to X_a .

In addition to the multiple activation steps shown in Fig. 4, processes such as blood coagulation appear to be regulated by specific protein inhibitors (19), which terminate the activation reactions once the product has been generated and restrict the site of blood coagulation to the area of the injury. In this case, a pulse of protease is generated only in the brief interval between zymogen activation and the subsequent formation of an inactive enzyme-inhibitor complex, as represented schematically in Fig. 5. This general scheme illustrates that the physiological signal is transduced to a chemical event by conversion of zymogen X to protease X_a , which in turn converts Y to Y_a and is then inactivated by the inhibitor I_x . In turn, Y_a exists long enough to activate Z and is then inhibited by I_y . Thus at each step, the duration and breadth of the cascade is controlled to some extent by the concentration of inhibitors, whereas the degree of amplification is determined by the number of steps between X_a and Z. The ultimate change in microenvironment is determined by the nature of the active protein that is finally expressed. There are only a few examples of cascading systems of this kind, but certainly now that the principle is established many more will be found. For instance, premature activation of the pancreatic enzymes shown in Fig. 3 is inhibited at the level of protease Y (trypsin) by pancreatic trypsin inhibitors. The various intermediate proteases of the blood coagulation system are inactivated by specific inhibitors, such as antithrombin III and α_2 -macroglobulin, whereas fibrinolysis (i.e., the dissolution of the fibrin clot) and the complement reaction each involve different sets of proteolytic activation reactions and different inhibitors such as α_1 -antitrypsin, α_2 -macroglobulin, C_1 -inactivator, and others (20, 21).

It seems important at this juncture to emphasize the fundamental difference between zymogens on the one hand and enzyme-inhibitor complexes on the other. Both appear inert by activity assays and both can be converted to the active enzyme under the appropriate circumstances. However, as in-

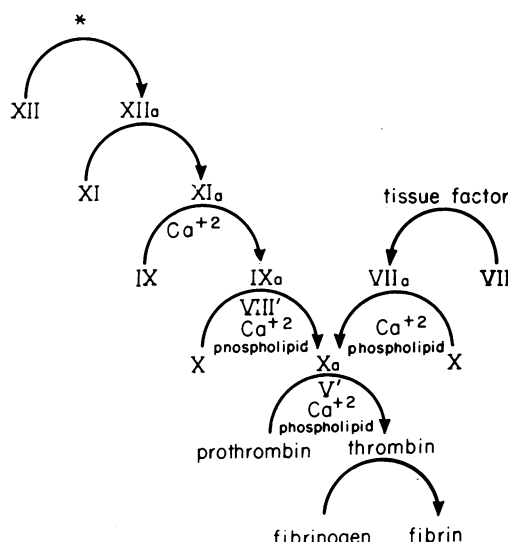


FIG. 4. An abbreviated representation of the cascade of consecutive zymogen activations involved in blood coagulation (3). The intrinsic pathway (left) and the extrinsic pathway (right) converge in activating factor X. The chemical nature of the initial stimulus (*) is not completely understood.

dicated schematically in Fig. 6, zymogens are inactive because the activation peptide confers an altered geometry on the molecule. The structure of the activation peptide is unrelated to that of the active site of the enzyme. On the other hand, macromolecular inhibitors contain regions that are complementary in shape to the active site of the enzyme and some have the characteristics of a pseudosubstrate (22, 23). In fact, only a few residues of the inhibitor molecule interact with the active site of the enzyme, whereas the bulk of the inhibitor serves mainly the function of a supporting structure. This is true of the pancreatic and soybean trypsin inhibitors, as well as of the carboxypeptidase inhibitor from potatoes that has been recently investigated in our laboratory (24).

Cascades of zymogen activation reactions do not necessarily operate in isolation but may influence one another in the initial stages, by positive or negative feedback regulation, thus adding another element of control. For instance, three plasma activation systems, the coagulation system, the fibrinolysis system, and the kallikrein system, interact with one another at key points as shown diagrammatically in Fig. 7. The activation of Hageman factor (factor XII) is enhanced by kallikrein, which in turn is generated from prekallikrein under the influence of activated Hageman factor (3, 25, 26). Kallikrein in turn cata-

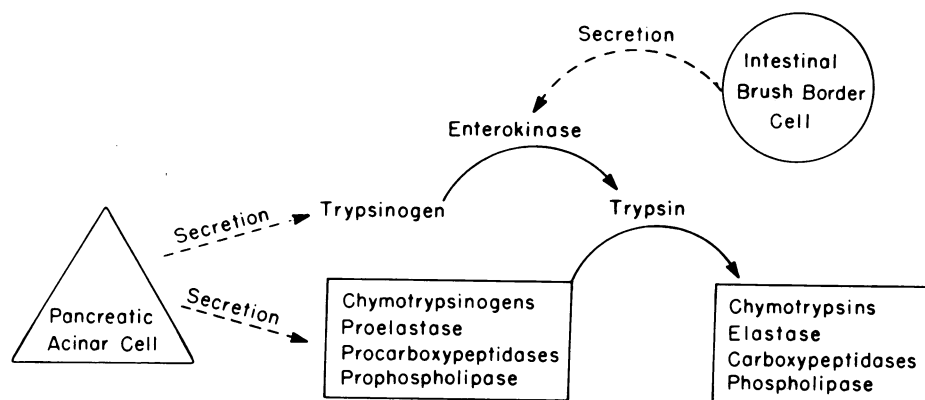


FIG. 3. Activation of zymogens in the duodenum involves a two-step cascade. Pancreatic trypsinogen is converted to trypsin by enterokinase. Trypsin in turn converts the other zymogens to active enzymes.

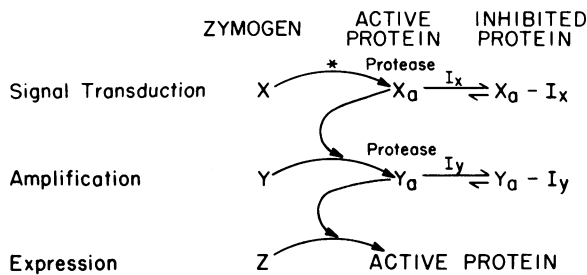


FIG. 5. Generalized scheme of control processes regulated by zymogens, proteases, and their inhibitors. Zymogens (X, Y, and Z) are activated sequentially in processes such as blood coagulation or complement activation. Protease intermediates (X_a, Y_a) may be inactivated by specific inhibitors (I_x, I_y) to limit their action. A physiological signal (*) initiates the cascade by converting zymogen X to protease X_a.

lyzes the formation of kinins (e.g., bradykinin) from their respective precursors, the kininogens (27). Activated Hageman factor is also believed to enhance the conversion of plasminogen proactivator to the activated form, though this process may involve a number of as yet unrecognized intermediates (26). Another focus of interaction of these systems is the degradation of fibrin by plasmin. Some of these reactions can also be inhibited by intermediates of the activation reactions as well as by specific plasma protease inhibitors (28, 29).

At present, descriptions of cascade processes deal primarily with plasma proteases, but fragmentary evidence is consistent with the occurrence of such complex processes elsewhere. For example, it is not yet clear whether the proteases that activate procollagen (8), proinsulin (5), prococoonase (30), and proacrosin (31) are directly triggered by specific enzymes or whether consecutive protease-mediated steps control these transformations. The sequence of discoveries in this field is the reverse of the physiological sequence of events because the starting point for investigation is usually the final product. Subsequent discovery of a precursor of this product leads to inquiries about a catalyst for this transformation; this in turn leads to questions about the origin of this catalyst, and so on. Ultimately one must identify for each process the mechanism that triggers the first step (compare Table 1).

Experimental investigations have usually focused on the physiologically active protein generated by zymogen activation (e.g., an enzyme, hormone, or self-assembling protein), whereas the peptide fragment has been generally ignored. However, recent studies of the activation of complement factor C₃ indicate (32) that a physiologically active peptide, anaphylatoxin (C_{3a}) is released from C₃ together with an active protease (C_{3b}). Analogously, complement factor C₅ and prothrombin each yield upon activation both an active protease and a pharmacologically active peptide (15, 32, 33). It may be suggested that other activation processes also produce dual physiological products in response to a single stimulus and that such reactions serve to express multiple messages in a concerted manner.

Activation reactions involving limited proteolysis have received added significance from the recent observations that secretory proteins, such as the pancreatic zymogens, contain an additional polypeptide extension on the amino side of the normal activation peptide (34–37). Semantically, these proteins are denoted by the prefix *pre-*, whereas zymogens are usually denoted by the prefix *pro-* or the suffix *-ogen* (e.g., *preproparathyroid hormone*, *pretrypsinogen*). These “prezymogens” are thought to be the first products of messenger translation and evidence indicates that after passage of the nascent proteins through the membrane of the endoplasmic reticulum the peptide extension is severed from the zymogen by limited

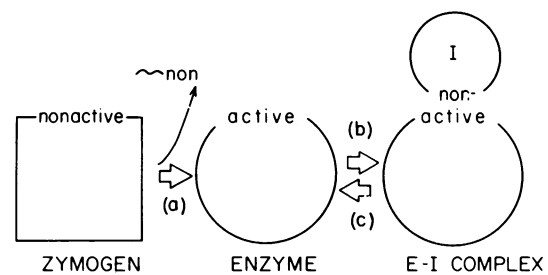


FIG. 6. Schematic illustration of the control of enzyme activity by (a) zymogen activation, (b) formation of an enzyme–inhibitor complex, and (c) dissociation of that complex. ~non represents the activation peptide.

proteolysis (34). While the precise significance of these events is currently under investigation, it seems certain that limited proteolysis, analogous to zymogen activation, transforms one precursor to another and that the mechanisms of conversion of zymogens to enzymes may serve as models for the transformation of prezymogen to zymogen.

EVOLUTION OF ZYMOGEN ACTIVATION

If limited proteolysis is indeed a significant physiological control mechanism, how did it come about during biological evolution? Did the primordial cells contain zymogens that subsequently became enzymes, or did they contain enzymes that later became inactivated by extension of the polypeptide chains so as to form zymogens? The available evidence seems to favor, but not prove, the latter alternative, which is also consistent with the empirical observation that relatively few zymogens have been found in prokaryotes or in less highly differentiated eukaryotic organisms.

It is a well-recognized fact that enzymes such as the mammalian serine proteases are a homologous set of proteins which operate by analogous catalytic mechanisms (38, 39). During evolution, those amino acid residues have been preserved that are essential for function, for the maintenance of conformation, or for both. In enzymes, these regions include, in particular, sequences around the active site, and in zymogens they might be expected to include regions around the site of activation. In the case of the serine proteases, the catalytic apparatus includes the so-called “charge-relay” system (40) formed by interaction of amino acid residues which are widely separated from each other in the linear sequence of the molecule. These include, in particular, aspartic acid 102, histidine 57, and serine 195. The alpha amino group of isoleucine 16 forms a salt linkage with aspartic acid 194. (The numbering system is that of bovine chymotrypsinogen A.) An abbreviated presentation of such patterns of homology among serine proteases is given in Fig. 8. The enzymes include three plasma proteases and their homology with the pancreatic serine proteases suggests that both families of enzymes have diverged from a common ancestor (41, 42).

Comparison of the zymogens of these proteases, on first sight, indicates that they are so different in size, composition, and sequence as to preclude a common ancestral relationship. For instance, in the case of bovine trypsinogen, the activation peptide is composed of six residues, whereas in the other extreme, the fragment released during the activation of prothrombin is larger than the enzyme thrombin itself. However, upon closer inspection, at least three classes of activation peptides seem evident. As shown in Fig. 9, trypsinogen activation peptides are homologous and related to the activation peptide released during the generation of the moth enzyme cocoonase from its precursor. Analogously, the activation peptides of

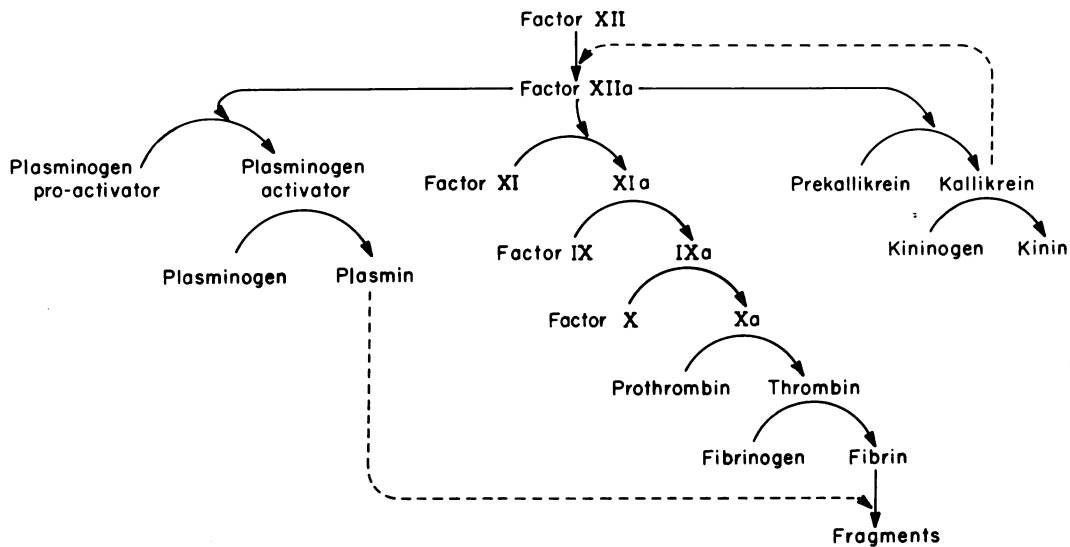


FIG. 7. Interactions among some plasma zymogens and proteases in the fibrinolytic, coagulation, and kinin systems (7, 25-27). For details, see the text.

various chymotrypsinogens (A, B, and C) and proelastase are homologous, although they appear to be quite different from those of the trypsinogen group. A third structural class of activation peptides comprises certain blood coagulation proteases, particularly those of prothrombin, factor X, and factor IX. The structural relationship among these zymogens is shown schematically in Fig. 10. Whereas prothrombin and factor IX are synthesized as single polypeptide chains and only subsequently during activation are converted into two-chain structures, factor X appears to consist even initially of two chains (7). The amino-terminal sequences of the single chains of prothrombin and factor IX and of the light chain of factor X show evidence of homology, but in the case of prothrombin and factor X no such similarity is evident in the sequences immediately preceding the peptide bond that is cleaved during the activation reaction (43, 44). If in fact the activation peptides of the three groups of serine proteases, i.e., the trypsinogens, the chymotrypsinogens, and the blood coagulation proteases, are unre-

lated, one would conclude that the homologous enzymes arose first and that the prefixes were subsequently added as independent events following gene duplication and divergence of functions (45). However, one cannot rule out the possibility that the activation peptides of the serine proteases represent distant homologous sequences which are the products of a rapidly mutating gene, as for instance in the case of the fibrinopeptides. de Haen *et al.* (46) have in fact argued for such a homologous relationship among these various zymogens.

The activation peptides of the three blood coagulation zymogens are also of interest because they illustrate that the removal of the activation peptide during zymogen activation may have additional consequences besides the generation of an active site. It has been recently demonstrated by several workers that the amino-terminal region of prothrombin, and probably also of factors X and IX, contains an unusual amino acid, γ -carboxyglutamic acid, which appears to form an effective site for chelating calcium ions (43, 47-49). Formation of this amino

PANCREATIC PROTEASES

	16	57	102	195
Bovine chymotrypsin A	Ile-Val- <u>Asn</u> -Gly- - - - -Ala-Ala-His-Cys- - - - -Asn-Asp-Ile- - - - -Gly-Asp-Ser-Gly-			
Bovine trypsin	Ile-Val-Gly-Gly- -Ala-Ala-His-Cys- -Asn-Asp-Ile- -Gly-Asp-Ser-Gly-			
Porcine trypsin	Ile-Val-Gly-Gly- -Ala-Ala-His-Cys- -Asn-Asp-Ile- -Gly-Asp-Ser-Gly-			
Dogfish trypsin	Ile-Val-Gly-Gly- -Ala-Ala-His-Cys- -Asn-Asp-Ile- -Gly-Asp-Ser-Gly-			
Porcine elastase	<u>Val</u> -Val-Gly-Gly- -Ala-Ala-His-Cys- - <u>Tyr</u> -Asp-Ile- -Gly-Asp-Ser-Gly-			

PLASMA PROTEASES (Bovine)

Thrombin	Ile-Val- <u>Glu</u> -Gly- -Ala-Ala-His-Cys- - <u>Arg</u> -Asp-Ile- -Gly-Asp-Ser-Gly-
Factor Xa	Ile-Val-Gly-Gly- -Ala-Ala-His-Cys- - <u>Phe</u> -Asp-Ile- -Gly-Asp-Ser-Gly-
Factor IXa	<u>Val</u> -Val-Gly-Gly- - - - - n.d. - - - - - - - - n.d. - - - - -Gly-Asp-Ser-Gly-
Plasmin (human)	<u>Val</u> -Val-Gly-Gly- -Ala-Ala-His-Cys- - - - - n.d. - - - - -Gly-Asp-Ser-Gly-

OTHER PROTEASES

Cocoonase	Ile-Val-Gly-Gly- - - - - n.d. - - - - - - - - n.d. - - - - -Gly-Asp-Ser-Gly-
-----------	--

FIG. 8. Similarities in amino acid sequences immediately adjacent to components of the active sites of various serine proteases (43, 44, 48, 70-75). The residue numbers are those of chymotrypsinogen A. Underlined residues differ from the majority shown for a particular position in the sequence. n.d. = not determined.

GROUP I				
Bovine trypsinogen		Val-Asp-Asp-Asp-Asp-Lys	Ile-	↓
Dogfish trypsinogen		Ala-Pro-Asp-Asp-Asp-Asp-Lys	Ile-	
Prococoonase	(5 residues)---	Arg-Thr-Gln-Asp-Asp-Gly-Gly-Lys	Ile-	
GROUP II				
Bovine chymotrypsinogen A	(7 res)---	Pro-Val-Leu-Ser-Gly-Leu-Ser-Arg	Ile-	
Bovine chymotrypsinogen B	(7 res)---	Pro-Val-Leu-Ser-Gly-Leu-Ala-Arg	Ile-	
Bovine chymotrypsinogen C	(5 res)---	Phe-Gln-Pro-Asn-Leu-Ser-Ala-Arg	Val-	
Lungfish proelastase A	(3 res)---	Pro-Ser-Tyr-Pro-Pro-Thr-Ala-Arg	Val-	
GROUP III				
Bovine prothrombin	Ala-Asn-Lys-Gly-Phe-Leu-Gla-Gla-----	(307 res)-----Phe-Glu-Ser-Tyr-Ile-Glu-Gly-Arg	Ile-	
Bovine Factor X	Ala-Asn-Ser- - -Phe-Leu-Gla-Gla--	(133 res) (43 res)--Glu-Asp-Gly-Ser-Gln-Val-Val-Arg	Ile-	
Bovine Factor IX	Tyr-Asn-Ser-Gly-Lys-Leu-Glx-Glx----	(ca. 147 res)-----	Val-	
Bovine plasminogen	Glu-Pro-Leu-Asp-Asp-Tyr-Val-Asn----	(ca. 550 res)-----	Val-	

FIG. 9. Homologous amino-terminal sequences of three groups of zymogens of serine proteases (30, 41, 43, 44, 46, 48, 70, 76). The vertical arrow denotes the site of limited proteolysis during zymogen activation. Gla indicates γ -carboxyglutamic acid.

acid is a prerequisite for the activation of prothrombin by factor X, and its absence in certain patients is associated with disorders of the blood coagulation process (49). The activation of prothrombin by factor X_a is accelerated by calcium ions, as is the activation of factor X by factor IX_a. Both activation reactions require, in addition, phospholipid and other coagulation proteins (factors V and VIII, respectively). Whereas activated factor X (X_a) stays bound to the phospholipid micelle, thrombin does not, because the calcium-binding site is released during the activation of prothrombin, but the binding site of factor X remains attached to this zymogen even after activation (3).

MECHANISM OF ZYMOGEN ACTIVATION

It is of foremost interest to inquire about the mechanism by which the cleavage of a single polypeptide bond in a zymogen generates enzymatic function. Does limited proteolysis merely remove an obstruction from the active site, or does it induce a conformational change that generates an active site? It is a well-accepted dogma that the amino acid sequence of a protein

provides the determinants of its three-dimensional structure and that chemical alterations of the backbone or the side chains could disrupt the balance of forces that dictate a particular conformation. A change in conformation could induce the *de novo* assembly of the catalytic apparatus, the generation of the substrate-binding site, or the removal of an obstruction from both (see Fig. 11). In the best-studied cases, i.e., the x-ray crystallographic comparison of the structures of chymotrypsinogen and chymotrypsin, the observed differences were so subtle that it was not possible to decide unambiguously which structural changes were crucial for zymogen activation and which were incidental (50). The recent discovery that many zymogens possess weak but intrinsic enzymatic activity has made it possible to examine the activation process by kinetic

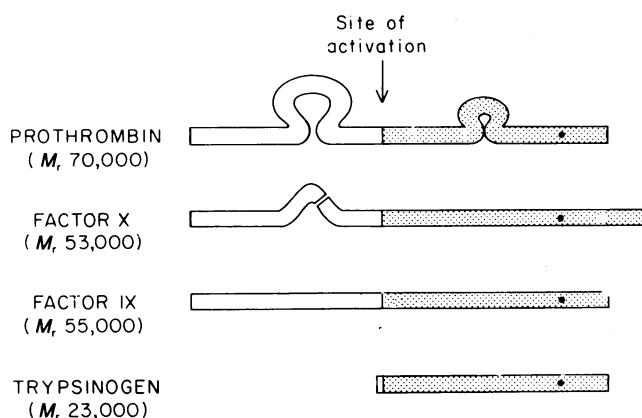


FIG. 10. Comparison of the lengths of the polypeptide chains of the zymogen forms of four homologous serine proteases. The stippled segments denote the homologous enzymes and the dots the reactive serines. M_r = molecular weight.

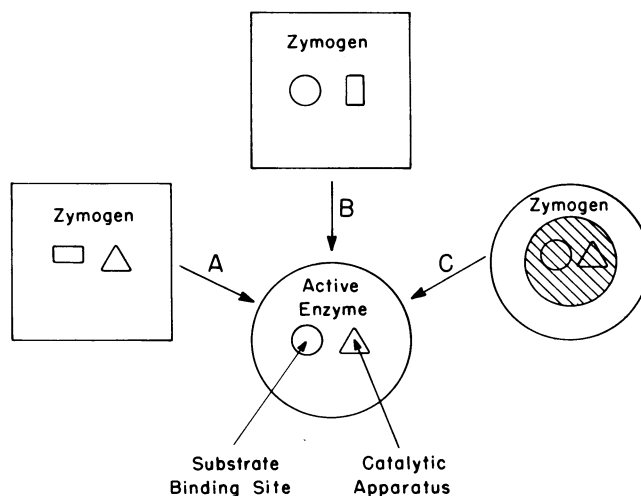


FIG. 11. Schematic illustration of three possible modes of zymogen activation. Pathway A: the conversion of zymogen to enzyme induces a conformational change that improves the substrate binding site. Pathway B: the binding site preexists in the zymogen and zymogen activation induces the formation of an effective catalytic site. Pathway C: the activation peptide (hatched) occludes the entire active site, which becomes exposed only after removal or dislocation of this peptide segment.

and spectral analysis of the zymogen before activation and of the enzyme afterwards (51–55). These studies, to be described briefly in the following paragraph, indicate that the catalytic apparatus of trypsin, chymotrypsin, and probably other zymogens of serine proteases, is largely preexistent in the zymogen forms and that during activation the effectiveness of the binding site is improved more than 1000-fold. The evidence can be briefly summarized as follows.

Several zymogens, like their respective enzymes, react stoichiometrically with active-site-directed inhibitors such as diisopropylphosphofluoridate to form stable covalent compounds at a reaction rate which is 10^4 to 10^5 times slower than that of enzyme (52). Certain other pseudo-substrates react not only with the enzyme but also with the parent zymogen. In both cases stable acyl-intermediates can be isolated at low pH and then deacylated at higher pH (54, 55). The second-order rate constants of acylation are 10^4 to 10^5 times lower for the zymogen than for the enzyme, but the deacylation rates differ only by factors ranging from 2 to 70. The pseudo-substrate methane sulfonyl fluoride, which has a low affinity for the enzyme, presumably because it is not bound to the substrate-binding pocket, also has a low affinity for the zymogen but the second-order rate constant differs not by 10^4 or 10^5 as in the case of diisopropylphosphofluoridate, but by only about 50-fold (56). Finally, a competitive inhibitor (*p*-aminobenzamidine) is bound by trypsinogen 10^3 to 10^4 times less firmly than by trypsin, presumably because the substrate-binding pocket is not fully developed prior to activation (54). All of these observations agree with the idea that in these zymogens the substrate-binding pocket is distorted but that the catalytic apparatus is largely preexistent.

It should be recognized, however, that the effectiveness of enzymes is determined by the geometries of both the enzyme and the substrate and that these have to be compatible to form the productive transition state during catalysis [in the case of serine proteases this appears to be a tetrahedral intermediate (57, 58)]. With ideal substrates, the geometries of enzyme and substrate are suitable to meet these requirements, whereas with poor substrates and pseudosubstrates the complex is less productive because of misalignment of susceptible bonds relative to the idealized state (M. A. Kerr, K. A. Walsh, and H. Neurath, manuscript in preparation). Zymogens may be relatively poor catalysts for a certain reaction because the geometry of the active site is unfavorable to bind substrates in a productive mode and to form the idealized transition state. Chymotrypsinogen and trypsinogen are inferior catalysts for all known substrates but activation by limited proteolysis changes the conformation of these zymogens, improves the binding of specific substrates, and also allows the transition state to be reached.

In a few cases, however, such conformational transitions occur seemingly without peptide bond cleavage. For instance, the conversion of human plasminogen to plasmin is induced by streptokinase, and as a first step the formation of a complex of streptokinase and plasminogen alone suffices to induce enzymatic activity without peptide bond cleavage (59, 60). Similarly, the activation of factor XII (Hageman factor) and of complement factor C_1 does not appear to involve peptide bond cleavage (4, 26). In the case of factor X, peptide bond cleavage is required for activation but the differences in enzymatic activity toward pseudosubstrates between zymogen and active enzyme are not nearly as great as in the case of the model system, chymotrypsinogen and chymotrypsin (M. A. Kerr, K. A. Walsh, and H. Neurath, manuscript in preparation). It thus appears that the various zymogens and their enzymes differ from each in their relative catalytic efficiency, depending upon the magnitude of the conformational changes required for

binding and aligning the substrate in an optimally productive mode.

CONCLUSIONS

While the activations of pancreatic zymogens such as trypsinogen and chymotrypsinogen are the best understood examples of the induction of biological activity by limited proteolysis, it is important to emphasize that they represent only examples of a great variety of physiological functions that are induced by limited peptide bond cleavage of precursor proteins. The product of activation may be an enzyme, a hormone, a pharmacologically active peptide, or a structural component of tissues, but the ultimate chemical event is limited proteolysis in every instance. Proteases can thus generate functions but they can also destroy them, and in this sense it is useful to think of limited proteolysis as a control element that can turn other reactions on and off by generating or destroying their catalysts. The on and off reactions are controlled by different switches, so to speak, because proteolysis is essentially an irreversible process and proteases are not endowed with repair functions. Zymogens are poised to respond to signals, to amplify them, and to respond to them irreversibly. Examples of such control processes are being increasingly observed and should be further thought of in the exploration of the control of metabolism and other physiological phenomena, including processes of development and differentiation.

This work has been supported in part by grants from the National Institutes of Health (GM15731) and the American Cancer Society (BC-91R).

1. Neurath, H. (1957) *Adv. Protein Chem.* **12**, 320–386.
2. Neurath, H. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 51–64.
3. Davie, E. W. & Fujikawa, K. (1975) *Annu. Rev. Biochem.* **44**, 799–829.
4. Müller-Eberhard, H. J. (1975) *Annu. Rev. Biochem.* **44**, 697–724.
5. Steiner, D. F., Kemmler, W., Tager, H. S., Rubenstein, A. H., Lernmark, A. & Zuhlke, H. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 531–549.
6. Zanefeld, L. J. D., Polakoski, K. L. & Schumacher, G. F. B. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 683–706.
7. Cabib, E. & Farkas, V. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2052–2056.
8. Bornstein, P. (1974) *J. Supramol. Struct.* **2**, 108–120.
9. Laemmli, U. K. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 661–676.
10. Linderstrøm-Lang, K. U. & Ottesen, M. (1949) *C. R. Trav. Lab. Carlsberg* **26**, 403–442.
11. Richards, F. M. & Vithayathil, P. J. (1959) *J. Biol. Chem.* **234**, 1459–1465.
12. Schimke, R. T. & Doyle, D. (1970) *Annu. Rev. Biochem.* **39**, 929–976.
13. Ottesen, M. (1967) *Annu. Rev. Biochem.* **36**, 55–76.
14. Stadtman, E. R. (1970) in *The Enzymes*, ed. Boyer, P. D. (Academic Press, New York), 3rd ed., Vol. 1, pp. 397–459.
15. Walsh, K. A. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 1–11.
16. Maroux, S., Baratti, J. & Desnuelle, P. (1971) *J. Biol. Chem.* **246**, 5031–5039.
17. Macfarlane, R. G. (1964) *Nature* **202**, 498–499.

18. Davie, E. W. & Ratnoff, O. D. (1964) *Science* **145**, 1310-1312.
19. Heimbürger, N. (1974) in *Bayer-Symposium V, Proteinase Inhibitors*, eds. Fritz, H., Tschesche, H., Greene, L. J. & Truscheit, E. (Springer-Verlag, New York), pp. 14-22.
20. Harpel, P. C., Mosesson, M. W. & Cooper, N. R. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 387-404.
21. Jeppsson, J. and Laurell, C. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 405-414.
22. Huber, R., Kikla, D., Steigemann, W., Deisenhofer, J. & Jones, A. in *Bayer Symposium V, Proteinase Inhibitors*, eds. Fritz, H., Tschesche, H., Green, L. J. & Truscheit, E. (Springer-Verlag, New York), pp. 497-512.
23. Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H. & Blow, D. M. (1974) *Biochemistry* **13**, 4212-4228.
24. Hass, G. M., Ako, H., Grahn, D. T. & Neurath, H. (1976) *Biochemistry* **15**, 93-100.
25. Nagasawa, S., Takahashi, H., Koida, M., Suzuki, T. & Schoenmakers, J. G. G. (1968) *Biochem. Biophys. Res. Commun.* **32**, 644-649.
26. Ulevitch, R. J., Cochrane, C. G., Revak, S. D., Morrison, D. C. & Johnston, A. R. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 85-93.
27. Pisano, J. J. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 199-222.
28. Astrup, T. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 343-355.
29. Dan, K. & Reich, E. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 357-366.
30. Felsted, R. L., Kramer, K. J., Law, J. H., Berger, E. & Kafatos, F. C. (1973) *J. Biol. Chem.* **248**, 3012-3020.
31. Meizel, S. & Mukerji, S. K. (1975) *Biol. Reprod.* **13**, 83-93.
32. Hugli, T. E. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 273-290.
33. Jesty, J. & Esnouf, M. P. (1973) *Biochem. J.* **131**, 791-799.
34. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67**, 835-862.
35. Devillers-Thiery, A., Kindt, T., Scheele, G. & Blobel, G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 5016-5020.
36. Kemper, B., Habener, J. F., Ernst, M. D., Potts, J. T., Jr. & Rich, A. (1976) *Biochemistry* **15**, 15-19.
37. Schechter, I. & Burstein, Y. (1976) *Biochem. Biophys. Res. Commun.* **68**, 489-496.
38. Neurath, H., Walsh, K. A. & Winter, W. P. (1967) *Science* **158**, 1638-1644.
39. Hartley, B. S. & Shotton, D. M. (1971) in *The Enzymes*, ed. Boyer, P. D. (Academic Press, New York), 3rd ed., Vol. 3, pp. 323-373.
40. Blow, D. M. (1976) *Acc. Chem. Res.* **9**, 145.
41. Enfield, D. L., Ericsson, L. H., Fujikawa, K., Titani, K., Walsh, K. A. & Neurath, H. (1974) *FEBS Lett.* **47**, 132-135.
42. Titani, K., Hermodson, M. A., Fujikawa, K., Ericsson, L. H., Walsh, K. A., Neurath, H. & Davie, E. W. (1972) *Biochemistry* **11**, 4899-4903.
43. Enfield, D. L., Ericsson, L. H., Walsh, K. A., Neurath, H. & Titani, K. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 16-19.
44. Titani, K., Fujikawa, K., Enfield, D. L., Ericsson, L. H., Walsh, K. A. & Neurath, H. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3082-3086.
45. Neurath, H., Walsh, K. A. & Gertler, A. (1974) in *Metabolic Interconversion of Enzymes, 1973*, eds. Fischer, E. H., Krebs, E. G., Neurath, H. & Stadtman, E. R. (Springer-Verlag, New York), pp. 301-312.
46. de Haen, C., Neurath, H. & Teller, D. C. (1975) *J. Mol. Biol.* **92**, 225-259.
47. Stenflo, J. (1974) *J. Biol. Chem.* **249**, 5527-5535.
48. Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Morris, H. R. & Dell, A. (1974) *FEBS Lett.* **44**, 189-193.
49. Stenflo, J., Fernlund, P. & Roepstorff, P. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 111-122.
50. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. & Xuong, N. H. (1970) *Biochemistry* **9**, 1997-2009.
51. Robinson, N. C., Neurath, H. & Walsh, K. A. (1973) *Biochemistry* **12**, 420-426.
52. Morgan, P. H., Robinson, N. C., Walsh, K. A. & Neurath, H. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3312-3316.
53. Kassell, B. & Kay, J. (1973) *Science* **180**, 1022-1027.
54. Gertler, A., Walsh, K. A. & Neurath, H. (1974) *Biochemistry* **13**, 1302-1310.
55. Kerr, M. A., Walsh, K. A. & Neurath, H. (1975) *Biochemistry* **14**, 5088-5094.
56. Morgan, P. H., Walsh, K. A. & Neurath, H. (1974) *FEBS Lett.* **41**, 108-110.
57. Henderson, R., Wright, C. S., Hess, G. P. & Blow, D. M. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 63-70.
58. Robertus, J. D., Kraut, J., Alden, R. A. & Birktoft, J. J. (1972) *Biochemistry* **11**, 4293-4303.
59. Nagendra, K., Reddy, N. & Markus, G. (1972) *J. Biol. Chem.* **247**, 1683-1691.
60. Summaria, L., Arzadon, L., Bernabe, P. & Robbins, K. C. (1974) *J. Biol. Chem.* **249**, 4760-4769.
61. Skeggs, L. T., Jr., Kahn, J. R., Lentz, K. & Shumway, N. P. (1957) *J. Exp. Med.* **106**, 439-453.
62. Tager, H. S. & Steiner, D. F. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2321-2325.
63. Holstein, T. J., Stowell, C. P., Quevedo, W. C., Jr., Zarcaro, R. M. & Bienieki, T. C. (1973) *Yale J. Biol. Med.* **48**, 560-571.
64. McGuire, J., Newman, D. & Barisas, G. (1973) *Yale J. Biol. Med.* **48**, 572-582.
65. Dohke, K. (1973) *Arch. Biochem. Biophys.* **157**, 203-209.
66. Leckie, B. (1973) *Clin. Sci.* **44**, 301-304.
67. de Haas, G. H., Slotboom, A. J., Bensen, P. P. M., Van Deenen, L. L. M., Maroux, S., Puigserver, A. & Desnuelle, P. (1970) *Biochim. Biophys. Acta* **221**, 31-53.
68. Harper, E., Bloch, K. J. & Gross, J. (1971) *Biochemistry* **10**, 3035-3041.
69. Kreil, G. (1973) *Eur. J. Biochem.* **33**, 558-566.
70. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C.), Vol. 5.
71. Hermodson, M. A., Ericsson, L. H., Neurath, H. & Walsh, K. A. (1973) *Biochemistry* **12**, 3146-3153.
72. Titani, K., Ericsson, L. H., Neurath, H. & Walsh, K. A. (1975) *Biochemistry* **14**, 1358-1366.
73. Robbins, K. C., Bernabe, P., Arzadon, L. & Summaria, L. (1972) *J. Biol. Chem.* **247**, 6757-6762.
74. Robbins, K. C., Bernabe, P., Arzadon, L. & Summaria, L. (1973) *J. Biol. Chem.* **248**, 1631-1633.
75. Kramer, K. J., Felsted, R. L. & Law, J. H. (1973) *J. Biol. Chem.* **248**, 3021-3028.
76. Wallen, P. & Wiman, B. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2, pp. 291-303.
77. Perlmann, G. E. & Lorand, L., eds. (1970) *Methods in Enzymology* (Academic Press, New York), Vol. 19.
78. Boyer, P. D., ed. (1971) *The Enzymes*, 3rd ed. (Academic Press, New York), Vol. 3.
79. Reich, E., Rifkin, D. & Shaw, E., eds. (1975) *Cold Spring Harbor Conferences on Cell Proliferation* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), Vol. 2.