NEW COMPREHENSIVE BIOCHEMISTRY VOLUME 16

general editors:
A. NEUBERGER & L.L.M. VAN DEENEN

# Hydrolytic Enzymes

edited by
A. Neuberger & K. Brocklehurst

ELSEVIER

# Hydrolytic Enzymes

*Editors*

A. NEUBERGER and K. BROCKLEHURST

*The Lister Institute of Preventive Medicine, Charing Cross Hospital Medical School, St. Dunstan's Rd., London W6 8RP (U.K.), and Department of Biochemistry, Medical College of St. Bartholomew's Hospital, University of London, Charterhouse Square, London EC1M 6BQ (U.K.)*


1987
ELSEVIER
AMSTERDAM · NEW YORK · OXFORD

ISBN for the series: 0-444-80303-3
ISBN for the volume: 0-444-80886-8

Published by:

Elsevier Science Publishers B.V.
(Biomedical Division)
P.O. Box 211
1000 AE Amsterdam
(The Netherlands)

Sole distributors for the U.S.A. and Canada:
Elsevier Science Publishing Company, Inc.
52 Vanderbilt Avenue
New York, NY 10017
(U.S.A.)

**Library of Congress Cataloging-in-Publication Data**
(Revised for volume 16)

New comprehensive biochemistry.

Vol. 16 published: Amsterdam; New York: Elsevier Science Publishers B.V. (Biomedical Division); New York, N.Y. U.S.A.: Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co.
Includes bibliographical references and indexes.
Contents: v. 1. Membrane structure / editors, J.B. Finean and R.H. Michell -- v. 3 Stereochemistry / editor, Ch. Tamm -- v. 4. Phospholipids / editors, J.N. Hawthorne and G.B. Ansell -- [etc.] -- v. 16. Hydrolytic enzymes / editors, A. Neuberger and K. Brocklehurst.
1. Biological chemistry--Collected works.
I. Neuberger, Albert. II. Deenen, Laurens L.M. van.
[DNLM 1. Membranes--Anatomy and histology. WI NE372F v.1 / QS 532.5.M534]
QD415.N48 574.19'2 81-3090
ISBN 0-444-80303-3 (Elsevier/North-Holland : set)

Printed in The Netherlands

# HYDROLYTIC ENZYMES

# New Comprehensive Biochemistry

Volume 16

*General Editors*

A. NEUBERGER
*London*

L.L.M. van DEENEN
*Utrecht*


ELSEVIER
AMSTERDAM · NEW YORK · OXFORD

# Preface

A volume dedicated to Hydrolytic Enzymes was perceived as useful for two reasons. In the first place, a number of these enzymes are not dealt with fully in those volumes of this series in which systems and events are discussed principally in a particular metabolic or physiological context. Secondly, it seems appropriate to bring together discussion of some of the enzymes that became the focus of attention in the 1960s when our understanding of the function of protein molecules was revolutionised by the application of X-ray crystallography. At that time, an account of the structure of myoglobin was rapidly followed by reports of the structures of lysozyme, carboxypeptidase, ribonuclease, chymotrypsin and papain, which permitted for the first time the results of mechanistic study by kinetic and protein chemical methods to be thought about within a realistic structural framework. Six of the eight chapters are devoted to various aspects of proteolysis. This emphasis is not inappropriate in view of the many advances in the search for chemical understanding of biological phenomena that were achieved during the study of proteolytic enzymes. They were among the first enzymes to be highly purified and crystallised and much of our understanding of molecular aspects of catalytic mechanism and specificity is founded on the study of these enzymes. Four of the chapters deal with the different types of proteinase that are differentiated on the basis of Hartley's idea of classifying proteinases by catalytic mechanism rather than by origin, specificity or physiological function. These chapters are complemented by one on proteinase inhibitors and by a short review of intracellular proteolysis. The latter includes a brief discussion of ATP-dependent proteolysis by ubiquitin, which will be extended in a subsequent volume dealing with protein metabolism. The final two chapters deal respectively with pancreatic ribonuclease A, the best characterised of the endoribonucleases, and with the phosphomonoesterases. A particular regret, in view of the central importance of lysozyme in the development of studies on structure, specificity and mechanism, is the unavoidable omission of a chapter on glycosidases. It is hoped that this omission will be rectified in a subsequent volume.

The major development of the 1960s in providing three-dimensional structures of enzymes at atomic resolution is being augmented in the 1980s by the application of DNA technology to provide designed structural variation in individual amino acid residues by site-directed mutagenesis. This approach should go some way towards obviating the largest single problem that has held back mechanistic study of enzyme catalysis, namely the inability to vary systematically the structure of both or all of the reactant molecules.

We wish to record our thanks to the authors both for their excellent contributions and for their helpful cooperation in the editorial process.

London
December 1987

A. Neuberger
K. Brocklehurst

# Contents

A. Neuberger and K. Brocklehurst (Eds.), *Hydrolytic Enzymes*




CHAPTER 1

# Aspartyl proteinases

JOSEPH S. FRUTON

*Yale University, New Haven, CT (U.S.A.)*

## 1. Introduction

### (a) Historical background

The aspartyl proteinases represent one of the four known main classes of enzymes that act at interior peptide bonds of proteins and oligopeptides (endopeptidases); the other classes are denoted serine proteinases, cysteine proteinases and metalloproteinases. Because of their optimal action at pH 1.5–5, the aspartyl proteinases were previously named acid proteinases. With the recognition that particular carboxyl groups in these enzymes are essential for catalysis, the term carboxyl proteinase was then used. The identification of these groups as belonging to aspartyl residues in several members of this class has led to the currently-preferred terminology. The term 'aspartyl proteinase' (aspartic proteinase and aspartate proteinase have also been used) is more appropriate than 'acid proteinase' because some enzymes now known to belong to this class act optimally on their substrates near pH 7.

Few enzymes occupy a more important place in the history of biochemistry than the one found in 1834 by Johann Nepomuk Eberle (1798–1834) in extracts of gastric mucosa. Two years later, Theodor Schwann (1810–1882) characterized this 'ferment', named it *pepsin*, and established its physiological role in the mammalian digestion of food proteins [1]. During the succeeding 60 years, pepsin was considered to be the prototype of the 'unorganized ferments' (Kühne named them *enzymes* in 1876) as distinct from the 'organized ferments' responsible for such processes as the fermentation of sugar by yeast [2]. Many efforts were made to purify pepsin; the work of Ernst Wilhelm von Brücke (1819–1892) and of Cornelis Adrianus Pekelharing (1848–1922) is especially noteworthy. The high point came in 1930, when John Howard Northrop (b. 1891) described the crystallization of pig pepsin [3]. Although this achievement followed the crystallization of urease by James Sumner, it was Northrop's massive evidence for the protein nature of pepsin that led to the rejection of the view, advocated

by Richard Willstätter during the 1920's, that enzymes are small catalytic molecules adsorbed on inactive protein carriers [4].

Another important discovery made in this field before 1900 was the observation by John Newport Langley (1852–1925) that a slightly alkaline extract of gastric mucosa contains a material (*pepsinogen*) which is converted to pepsin on acidification of the extract [5]. The crystallization of pig pepsinogen in 1938 by Roger Moss Herriott (b. 1908) made possible the incisive study of its conversion to pepsin [6]. The work of Northrop and Herriott thus marks the beginning of the modern study of pepsin as a protein and as a catalytic agent, and has influenced the investigation of the other enzymes now considered to be aspartyl proteinases.

*(b) Occurrence and nomenclature*

In all vertebrates the gastric juice contains one or more pepsins that arise from secreted pepsinogens; the latter are produced mainly in the chief cells (zymogen cells) of the fundus (corpus) [7,8]. The secretion of the pepsinogens is under vagal control both directly on the oxyntic glands of the fundus and indirectly through the release of peptide hormones (gastrins) from the pyloric glands [9].

Multiple forms of pepsinogen, and of the pepsins derived from them have been found in many vertebrates (e.g., man, monkey, pig, beef, rat, chicken, dogfish). Chromatographic separation of the components has shown that the predominant pepsin A (usually denoted pepsin) of adult mammals is accompanied by pepsin C (the currently-preferred name is *gastricsin*), as well as by the minor components denoted pepsin B and pepsin D [10]. Some investigators have given the individual chromatographic components Roman numerals, while others have numbered the gastric proteinases and their zymogens in the order of decreasing electrophoretic mobility at pH 5.0 or 8.5 respectively [11]. Immunochemical methods have also been applied to the differentiation and numbering of human gastric proteinases [12]. In the gastric juice of fetal and newborn mammals the major pepsin-like enzyme is *chymosin* (derived from the zymogen prochymosin); this was its original name [13], but for many years it was called *rennin* because it is the chief enzymic component of rennet, the calf-stomach (abomasus) extract used in the manufacture of cheese [14]. Regrettably, the nomenclature of the mammalian gastric proteinases has been in a confused state because of differences in the terminology used by various groups of investigators; for a helpful clarification, see Foltmann [15].

Among the aspartyl proteinases from vertebrates is the kidney enzyme *renin* (also present in submaxillary tissue), whose important physiological function is the formation from plasma angiotensinogen of the decapeptide angiotensin I, which is in turn cleaved by a 'converting enzyme' to the highly active pressor octapeptide angiotensin II [16,17]. It should be noted that although renin is now known to be an aspartyl proteinase, it is not an acid proteinase, since the pH optimum for its action is 6–8. Another aspartyl proteinase is the lysosomal enzyme *cathepsin D* present in many animal tissues (spleen [18], liver [19], uterus [20,21], thyroid [22], skeletal muscle [23], anterior pituitary [24], brain [25], seminal tissue [26], erythrocytes [27], lymphoid tissue [28]). A cathepsin D

and its inactive precursor from monkey lung appears to resemble gastricsin and its zymogen [29]. Among the aspartyl proteinases in plants are those present in Lotus seed [30] and in the insectivorous plants Nepenthes and Drosera [31].

There has been considerable interest in microbial acid proteinases in part because of a search for suitable rennet substrates. The enzymes subjected to the most intensive study have been penicillopepsin (from *Penicillium janthinellum*) [32], Rhizopus-pepsin (from *Rhizopus chinensis*) [33], and the acid proteinases from *Endothia parasitica* [34] and *Mucor miehei* [35]. Well-characterized microbial acid proteinases have also been isolated, and in some cases crystallized, from strains of *Acrocylindricum* sp. [36], *Aspergillus saitoi* [37], *Candida albicans* [38], *Cladosporium* sp. [39], *Fusarium moniliforme* [40], *Monascus kaoling* [41], *Mucor pusillus* [42], *Paecilomyces varioti* [43], *Penicillium duponti* [44], *Rhodotorula glutinis* [45], *Russula decolorans* [46], *Tramestes sanguinea* [47], *Trichoderma viride* [48], and both *Saccharomyces cerevisiae* and *S. carlsbergensis* (yeast proteinase A) [49]. No evidence is available for the existence of zymogens for the above microbial proteinases. The enzymes from *Endothia parasitica*, *Mucor miehei* and *Mucor pusillus* have been used in cheese manufacture.

These various aspartyl proteinases have in common the property of cleaving proteins (e.g., denatured hemoglobin, serum albumin) and suitable oligopeptides at pH 1.5–5.5. A widely-used diagnostic test is their inhibition by the naturally-occurring peptide pepstatin and by active-site-directed diazo compounds; these properties will be discussed later in this chapter. Although future study of other enzymes may show them to belong to the aspartyl proteinase family, in what follows primary attention will be given to those known to exhibit these properties. Among the acid proteinases that do not appear to be inhibited by pepstatin or diazo compounds is the one from *Scytalidium lignicolumn* [50].

Apart from the confusion in the naming of the mammalian gastric proteinases, mentioned previously, the nomenclature of the acid proteinases has not been aided by the Commission on Enzymes of the International Union of Biochemistry [51]. Proteolytic enzymes belonging to different classes have been given the same name and distinguished from each other only by the addition of a different capital letter and the assignment of different numbers. Moreover, the Commission has retained the long-outworn distinction between hydrolases and transferases for enzymes that act on peptide, ester and glycosidic bonds presented in the first edition of the treatise of Dixon and Webb [52].

### *(c) Purification*

In his work on crystalline pig pepsin, Northrop noted that the preparations differed considerably in homogeneity, as indicated by measurement of their solubility behavior. Part of the inhomogeneity was a consequence of the presence of peptide material, formed by autodigestion. Subsequently, Steinhardt performed a careful study of the solubility properties of crystalline pig pepsin, and gave clear evidence of its inhomogeneity as a protein [53]. After the introduction of ion-exchange chromatography for the fractionation of proteins, Ryle and his associates established the presence of the

minor components mentioned previously [10]. Additional factors that may contribute to heterogeneity are the presence of multiple gene products (made evident by amino acid substitutions) and different degrees of phosphorylation [54] or glycosylation [55].

At present, the preferred method for the preparation of apparently homogeneous pig pepsin A is rapid activation of crystalline pepsinogen (shown to be homogeneous by several criteria [56]), and passage of the mixture first through sulfoethyl Sephadex C-25 to remove peptides and then through Sephadex G-25 to remove salts. This pepsin preparation is homogeneous on hydroxylapatite or DEAE-cellulose, which may also be used to effect the purification of commercial preparations of crystalline pig pepsin [57]. Similar chromatographic procedures have been used for other aspartyl proteinases.

Some aspartyl proteinases have been purified by means of affinity chromatography on columns of Sepharose 4B or agarose to which a substrate analogue (e.g., L-Phe-D-Phe) [58] or pepstatin [59] has been attached by means of aminohexoyl bridges; hemoglobin-agarose columns have also been employed [60]. The purification of cathepsin D from various animal tissues has been attended with difficulty, as is indicated by the heterogeneity of the enzyme preparations that have been described [61–64].

Crystallization has been usually effected by means of ammonium sulfate or acetone. Aside from pig pepsin A, the following aspartyl proteinases have been obtained in crystalline form: the pepsins from beef [65] and salmon [66], calf chymosin [67], penicillopepsin [32], Rhizopus-pepsin [33], and the proteinases from *Endothia parasitica* [34], *Aspergillus saitoi* [37], *Mucor pusillus* [42], *Paecilomycetes varioti* [43], *Penicillium duponti* [44], *Rhodotorula glutinis* [45], and *Trametes sanguinea* [47]. The crystallization of chicken liver cathepsin D has been reported [68].

*(d) Assay*

A widely-used method is that introduced by Anson [69]; later investigators have made slight modifications in the procedure. Acid-denatured hemoglobin is the substrate at pH 1.8 and 37°C, and the release of cleavage products that are soluble in 3% trichloroacetic acid is measured spectrophotometrically at 280 nm. One unit of pepsin activity is usually defined as the amount of enzyme that produces an increase in absorbance of 0.001 per minute under the conditions of the assay. Commercial preparations of crystalline pig pepsin generally contain 2500–3000 units per milligram; the material obtained by chromatographic purification of pepsin produced by the rapid activation of pepsinogen assays at about 4000 units per milligram [57]. Other assay methods with protein substrates include the use of bovine serum albumin in place of hemoglobin. For chymosin and related acid proteinases, a useful assay procedure is the measurement of the rate of clotting of 10% reconstituted skim milk powder in the presence of $CaCl_2$ [14]. Some microbial acid proteinases catalyze the activation of trypsinogen at pH 3.4; the assay method introduced by Kunitz [70] has been modified [32].

Various synthetic peptides have been used in the assay of the aspartyl proteinases.

Among these substrates are compounds of the type A-Phe($NO_2$)-Y-B* (where Y = Phe, Leu etc.); the rate of cleavage of the Phe($NO_2$)-Y bond may be followed spectrophotometrically at 310 nm [71]. Older methods have involved the use of substrates of the type A-Phe-Y-OH (where Y = Tyr($I_2$), Phe etc.), and measurement of the rate of hydrolysis of the Phe-Y bond by the ninhydrin method [72]; this procedure has been automated [73].

## 2. *Molecular properties*

### *(a) Physical-chemical properties*

By means of the sedimentation-equilibrium method, values of 32700 ± 1200 and 40400 ± 1600 were obtained for the molecular weight of pig pepsin A and pepsinogen A respectively [74]. These values may be compared to 34644 and 39637 calculated from the amino acid sequences. Other methods (for example, sedimentation-velocity-diffusion, light scattering, osmotic pressure) gave values for pepsin ranging from 32000 to 35000. Estimates of the molecular weight of other gastric proteinases and of microbial acid proteinases (in some cases determined by means of sodium dodecyl sulfate–agar gel electrophoresis) have given values ranging from about 31000 to about 40000; for the zymogens the values range between 36000 and 43000 [75].

The aspartyl proteinases are acidic proteins, as a consequence of the preponderance of dicarboxylic acid residues as compared to the basic amino acid residues. In the case of pig pepsin, the paucity of lysine and arginine residues is especially marked. Early studies by Tiselius and Herriott suggested that the isoelectric point of pig pepsin lies below 1, since the protein still migrated as an anion at this pH value. This conclusion is probably incorrect, as is suggested by more recent studies in which the isoelectric focussing technique was employed [76–78]. However, in view of the heterogeneity of the preparations, and the extended time required in this method, no definite isoelectric point can be assigned, except to infer that pepsin A has a p*I* between 2 and 3. In the case of other aspartyl proteinases where the balance between acidic and basic amino acids is less extreme, the isoelectric points are between 3 and 5. In contrast to the extremely low isoelectric point of pig pepsin, that of pig pepsinogen is about 3.7; this difference is consistent with the cationic character of the peptide removed from the zymogen upon its conversion to pepsin (see Section 3(a)).

---

* Abbreviations (in alphabetical order) used in this chapter and not defined in the Recommendations of the IUPAC-IUB Joint Commission on Nomenclature on the Nomenclature and Symbolism for Amino Acids and Peptides [Eur. J. Biochem. (1984) 138, 9–37]: DAN, diazoacetyl-DL-norleucine methyl ester; Dns, 5-dimethylamino-1-naphthalenesulfonyl; EPNP, 1,2-epoxy-3-(*p*-nitrophenoxy)propane; Mns, 6-(*N*-methylanilino)-2-naphthalenesulfonyl; Nle, L-norleucyl; Nva, L-norvalyl; OP4P, 3-(4-pyridyl)propyl-1-oxy; Phe($NO_2$), *p*-nitro-L-phenylalanyl; Pla, β-phenyl-L-lactyl; Pol, L-phenylalaninol; TPDM, *p*-toluenesulfonyl-L-phenylalanyldiazomethane; TNS, 2-*p*-toluidinylnaphthalene-6-sulfonate; Tyr($I_2$), 3,5-diiodo-L-tyrosyl. The abbreviated designation of amino acid residues denotes the L form, except where otherwise indicated.