# THE PRIMARY STRUCTURE OF ATROPINESTERASE from Pseudomonas putida

*Atropa Belladonna*

J.G.M. HESSING

1.

De opvatting van Clement dat bij muizen die vergiftigd zijn met Soman, het aan aliesterase gebonden organofosfaat kan functioneren als een "Soman-depot" is in hoge mate onwaarschijnlijk.

Clement, J.G., (1982) Biochem. Pharmacology, 31, 4085-4088.

2.

De bewering dat de substraat specificiteit van alle serine-proteasen wordt bepaald door de aard van het aminozuur gelegen 6 residuen vóór de aktieve serine is niet algemeen toepasbaar.

Mole, J.E., and Niemann, M.A., (1980) J.of Biol.Chem. 255, 8472-8476

3.

De conclusie van Hovanec et al. betreffende het mechanisme van de veroudering van acetylcholinesterase dat geremd is met sarin, is gebaseerd op een ongefundeerde interpretatie van hun onderzoek naar de spontane reaktivering van het geremde enzym.

Hovanec, J.W., Broomfield, C.A., Steinberg, G.M., Lanks, K.W., and Lieske, C.N., (1977) Biochim.Biophys.Acta 483, 312-319

4.

Bij het bestuderen van het herstel van schade in DNA, die is aangebracht door blootstelling aan straling of chemische agentia, in zoogdiercellen waarvan -zoals te doen gebruikelijk is- het DNA tevoren is gemerkt door inbouw van radioaktieve bouwstenen, dient men er rekening mee te houden dat een deel van de effecten het gevolg kan zijn van de pre-labeling van de cellen.

Hirschi, M., Netrawali, M.S., Remsen, J.F., and Cerutti, P.A., (1981) Cancer Research 41, 2003-2007.

5.

De conclusie van Keesey en Demoss dat zij het trp-1 gen van Neurospora crassa hebben gecloneerd wordt niet volledig ondersteund door experimentele gegevens.

Keesey, J.K. JR., and Demoss, J.A., (1982) J. of Bacteriology 152, 954-958

6.

Het hanteren van biologische aktiviteit als criterium voor het vinden van de intacte genetische informatie van interferon genen is onjuist.
Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G., and Fiers, W., (1982) The EMBO Journal, 1, 953-958.

7.

De experimentele resultaten van Sage et al. geven geen steun aan hun conclusie dat specifieke detectie van 2-acetylaminofluorene-DNA adducten in celkernen heeft plaats gevonden, zodat hun conclusies omtrent inductie en herstel van deze DNA beschadigingen ongegrond zijn.
Sage, E., Gabelman, N., Mendez, F., and Bases, R., (1981) Cancer Letters 14, 193-204.

8.

De plaats waar onophoudelijk op hoog niveau chemie wordt bedreven is het levend organisme.

9.

De codering van genetische eigenschappen wordt beter bewaard door een computer dan door een levend organisme.

10.

Er moet veel tekst worden verwerkt, voordat een tekstverwerker een tekst kan verwerken.

Stellingen behorende bij het proefschrift:
The primary structure of atropinesterase from Pseudomonas putida.

J.G.M.Hessing
Leiden, 28 april 1983

# THE PRIMARY STRUCTURE OF ATROPINESTERASE from Pseudomonas putida

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE WISKUNDE EN NATUURWETENSCHAPPEN AAN DE RIJKSUNIVERSITEIT TE LEIDEN, OP GEZAG VAN DE RECTOR MAGNIFICUS DR. A.A.H. KASSENAAR, HOOGLERAAR IN DE FACULTEIT DER GENEESKUNDE, VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN TE VERDEDIGEN OP DONDERDAG 28 ARPIL 1983 TE KLOKKE 15.15 UUR

DOOR

JOHANNA GERARDINA MARIA HESSING

GEBOREN TE VOORBURG IN 1943

1983
DRUKKERIJ J.H. PASMANS B.V., 's-GRAVENHAGE

| | |
|---|---|
| Promotor: | Prof. Dr. Ir. A. Rörsch |
| Co-promotor: | Dr. F. Berends |
| Referenten: | Prof. Dr. W.J.H.M. Möller |
| | Prof. Dr. P.H. v. Knippenberg |

Overige leden van de promotiecommissie:

Prof. Dr. P. v.d. Putte
Prof. Dr. A.P. Zuur

Aan mijn moeder

CONTENTS

## LIST OF ABBREVIATIONS

### One- and three-letter amino acid abbreviations

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A | - | Ala | - alanine | M | - | Met | - methionine |
| C | - | Cys | - cysteine | N | - | Asn | - asparagine |
| D | - | Asp | - aspartic acid | P | - | Pro | - proline |
| E | - | Glu | - glutamic acid | Q | - | Gln | - glutamine |
| F | - | Phe | - phenylalanine | R | - | Arg | - arginine |
| G | - | Gly | - glycine | S | - | Ser | - serine |
| H | - | His | - histidine | T | - | Thr | - threonine |
| I | - | Ile | - isoleucine | V | - | Val | - valine |
| K | - | Lys | - lysine | W | - | Trp | - tryptophan |
| L | - | Leu | - leucine | Y | - | Tyr | - tyrosine |

| | |
|---|---|
| CMCys | - carboxymethylcysteine |
| Hse | - homoserine |
| Asx | - asparagine or aspartic acid |
| Glx | - glutamine or glutamic acid |

### General abbreviations

| | |
|---|---|
| A280 | - absorbance at 280 nm |
| A570 | - absorbance at 570 nm |
| BSA | - bovine serumalbumine |
| CNBr | - cyanogen bromide |
| CPA | - carboxypeptidase A |
| CPB | - carboxypeptidase B |
| DEAE | - diethylaminoethyl |
| DFP | - diisopropyl phosphorofluoridate |
| DIP | - diisopropyl phosphate |
| DNS | - dansyl (5-dimethylaminonaphthalene-1-sulfonyl) |
| DTE | - dithioerythritol |
| HPLC | - high-performance liquid chromatography |
| HVE | - high-voltage paperelectrophoresis |
| IEF | - isoelectric focusing |
| LAP | - leucinaminopeptidase |
| LSC | - liquid scintillation counting |
| NBS | - N-Bromosuccinimide |
| PA | - peptide analyzer |
| PAGE | - polyacrylamide gelelectrophoresis |
| PITC | - phenyl isothiocyanate |
| PTH | - phenylthiohydantoin |
| SA | - specific activity |
| SDS | - sodium dodecyl sulfate |
| sf | - superfine grade |
| Soman | - (1,2,2-trimethylpropyl methylphosphonofluoridate) |
| S.protease | - staphylococcal protease |
| TFA | - trifluoroacetic acid |
| TGA | - thioglycolic acid |
| TLC | - thin layer chromatography |
| TPCK | - L-(tosylamido-2-phenyl)ethyl chloromethyl ketone |
| UV | - ultraviolet |

## CHAPTER I INTRODUCTION

The study of atropinesterase (E.C.3.1.1.10) was based on the ever existing interest at the Medical Biological Laboratory to investigate the toxic action of organophosphorus compounds on living organism. Among these compounds are many insecticides like paraoxon and potential chemical warfare agents such as Sarin, Soman and Tabun. It has been shown that the toxic action of these compounds is predominantly caused by the inactivation of acetylcholinesterase, an enzyme possessing a key position in the cholinergic nervous system, where it is responsible for the rapid hydrolysis of the transmitter substance acetylcholine in the synapses. Inactivation of the esterase results in local accumulation of this neurotransmitter which interferes with proper transmission of stimuli.

Acetylcholinesterase (acetyl hydrolase E.C.3.1.1.7) belongs to the so-called serine-hydrolases, a group of hydrolytic enzymes with a closely related working mechanism. The name serine-hydrolase has been given to these enzymes because all possess one specific serine residue with an abnormally reactive hydroxylgroup in its side chain, which is an essential element of the catalytic center where substrate hydrolysis takes place (98). A common feature of all enzymes belonging to this category is the inhibition by organophosphates of the above-mentioned type, which is caused by the exclusive phosphorylation of this particular serine residue, in a reaction considered to be strictly analogous to the first step (acylation) in normal substrate hydrolysis (96). In fact, inhibition by organophosphates is a reliable criterium for the recognition of a serine-hydrolase. Beside acetylcholinesterase and other cholinesterases, a variety of hydrolytic enzymes, esterases and proteases, from animals, plants and micro-organisms belong to this category. Some of these serine-hydrolases are true esterases, they only possess esterase-activity, whereas others are proteases but are also capable to hydrolyse ester bonds.

One might obtain a better understanding of the working mechanism of serine-hydrolases and insight in the related interaction between

organophosphorus compounds and the active center of these enzymes, by performing structural studies, since it has been shown that the function of an enzyme and its reaction with different compounds is determined by its unique three-dimensional structure. All information to obtain the proper conformation is determined by the amino acid sequence of the protein, as was shown for the first time by Anfinsen (8). His findings appear to be generally true for non-processed proteins and in some cases, as he found for ribonuclease, for native proteins. Therefore, elucidation of the amino acid sequence of proteins seems to be an essential step in the determination of the secondary and tertiary structure. The amino acid sequence of a number of serine-enzymes has already been established (111). This primary sequence work has been carried out on serine-hydrolases that primarily function as proteolytic enzymes. Of some of these enzymes, such as chymotrypsin, trypsin and subtilisin, detailed structural information was obtained. All data together indicated that most serine-proteases should be classified in either of the two superfamilies of evolutionary related serine-hydrolases identified to date, the prothrombin-related enzymes and the subtilisins. According to present view, the mammalian proteases trypsin and chymotrypsin, which have distinctly differing specificities, have arisen by divergent evolution from one common ancestral form, whereas the protease subtilisin, produced by Bacillus subtilis, stems from a completely different ancestral sequence, but has converged to attain an almost identical catalytic structure at the active site (34), viz. the "charge relay" system. This system comprises the active serine, the imidazole side chain of a histidine residue and the β-carboxyl group of a buried aspartic acid residue, as was revealed firstly for chymotrypsin (14) and subsequently for other serine-proteases. Owing to the particular spatial positions of these groups and their mutual interactions, the hydroxylgroup of the serine is donated with an abnormally high reactivity as a nucleophile.
Contrary to the serine-enzymes mentioned above, the enzyme of our interest, acetylcholinesterase, does not possess proteolytic activity;

it is only known to function as an ester-hydrolase. This difference in specificity could implicate structural differences. Unfortunately, acetylcholinesterase is poorly accessible for biochemical or biophysical research, due to isolation and purification problems, the small amounts present in living organisms and its high molecular weight, i.e. 4 subunits of ca. 90,000 daltons. Other true esterases known to belong to the serine-hydrolases, such as serum-cholinesterase or liver-aliesterase have this poor accessibility in common with acetylcholinesterase. Therefore, Rörsch and Berends searched for a serine-enzyme that could replace acetylcholinesterase as a real esterase, but did not possess its unfavourable properties. For various reasons they focused their attention on bacterial atropinesterases (108) and found the enzyme produced by a certain strain of pseudomonas bacteria to be a good model-enzyme for the intended study of the esterase structure. The bacterial strain was one of nine atropinesterase-producing strains which were isolated by Rörsch and Berends from soil collected from between the roots of Atropa belladonna plants. The atropinesterase used in this study was isolated from Pseudomonas putida biotype A (strain L, PMBL1*) (108). The enzyme produced by the bacteria functions as an esterase during the first step in the metabolic pathway of atropine (126); it shows a specific interaction with this alkaloid and effectively catalyses the hydrolysis of atropine into tropine and tropic acid. The enzyme was characterized as a serine-hydrolase because it was inhibited by organophosphorus compounds such as Soman and DFP (108, 97). Like acetylcholinesterase the enzyme's natural function appears to be an esterase; up to now no indications have been obtained for the presence of proteolytic activity. Therefore, atropinesterase apparently belongs to the category of serine-esterases. These characteristics, and also the stability of the enzyme, the relatively low molecular weight (2 identical subunits of ca. 30,000) and the

* This number refers to the culture collection of the Medical Biological Laboratory TNO.

possibility to purify the enzyme and isolate it in sufficient yield, made atropinesterase a good model-enzyme to study the relationship between structure and function of the serine-esterases.

The principal aim of our study was to elucidate the primary structure of this atropinesterase to see whether this serine-esterase would be related to the previously studied proteases. To elucidate the primary structure the following strategy was used: large quantities of pseudomonas bacteria (strain L, PMBL1) were grown in order to isolate the pure, fairly stable enzyme in the milligram quantities necessary for biochemical and biophysical studies. For structural studies the enzyme was extensively purified to a purity of over 90% (Oosterbaan, to be published). One monomer appeared to consist of one polypeptide chain comprising one active serine residue. According to the amino acid composition per active center -determined from organophosphate binding- this chain appeared to be composed of about 260 amino acid residues.

Since a priori no homology with other enzymes could be expected which might help in the interpretation of the results, the primary structure of atropinesterase had to be determined with great reliability. To achieve this goal we used two indepent methods to degrade the protein, i.e. digestion with trypsin and cleavage with cyanogen bromide (CNBr). These methods were chosen because of the claimed specificity of the agents used. Degradation with CNBr, which cleaves peptide bonds at the C-terminal side of methionine (Met)-residues (59) should produce a set of 9 fragments because of the 8 Met residues present in the protein. Digestion with trypsin, which catalyzes the hydrolysis of the peptide bond at the C-terminal side of lysine (Lys) or arginine (Arg) (121) would result in the production of 29 fragments because of the presence of 21 Lys- and 7 Arg-residues. The number of fragments could be markedly reduced when -before digestion- the protein was treated with maleic anhydride (23) which blocks the α-aminogroup of the protein and the ε-aminogroup of Lys-residues; this blocking should restrict digestion to bonds involving Arg-residues only, resulting in the production of 8 fragments. Tryptic digestion of maleylated atropines-

terase and CNBr-cleavage of the protein should both produce fairly large peptides with an average length of about 30 amino acids provided that the residues are not too irregularly distributed over the total chain. The limited number of fragments produced might allow complete separation with the known fractionation methods and reduce the chance to miss a small fragment without recognition. Elucidation of the primary structures of all peptides obtained with both methods and comparison of the two sets of results should provide a reliable outcome for the complete sequence, as the result of the one method should confirm that of the other. Furthermore, in this way all information would have been obtained that is necessary to determine the order of the produced fragments in the total sequence.

The amino acid sequence in the peptides produced by the primary cleavage methods was established with the dansyl-Edman degradation procedure (57) or analogous methods. Usually the fragments were too large for complete sequence determination. In those cases secondary cleavage procedures were applied, with enzymes such as thermolysin or staphylococcal protease and/or chemical reagents, e.g. N-Bromosuccinimide (NBS).

In practice the strategy was not completely successful because of many difficulties encountered in producing the constituting fragments and in their isolation and purification, while in some cases problems arose during amino acid sequence determination. Application of the first method (tryptic digestion) produced a complete set of peptides, whereas the CNBr-method was hampered by separation problems, with the result that a few large fragments could not be isolated in pure state and could not be structured completely. Therefore, to put all peptides obtained with trypsin in the right position in the protein, additional approaches had to be applied, including degradation of atropinesterase with pepsin. Application of the methods mentioned above resulted in the elucidation of the complete amino acid sequence of atropinesterase. This enzyme now is the first serine-esterase the primary structure of which has been established.

Once the primary structure was elucidated it was investigated whether or not there existed any relationship between atropinesterase and the well-studied serine-proteases, which was the second aim of this study. To this end the amino acid sequence of atropinesterase was compared with the amino acid sequences of known serine-enzymes to determine any existing sequence homology. The enzyme was compared with two members of the superfamily of prothrombin-related enzymes: bovine trypsinogen and α-lytic protease from Myxobacter 495 and with one member of the subtilisin superfamily, subtilisin from Bacillus amyloliquiefaciens (BPN'). Comparison has also been made with two serine-enzymes from E.coli, viz. alkaline phosphatase and cephalosporinase. The last two enzymes are not related to the serine-proteases. They are not inhibited by organophosphorus compounds, but do contain an active serine residue. Although no striking difference in chain length was found between atropinesterase and the members of the superfamilies of subtilisins and prothrombin-related enzymes, the position of the active serine differed markedly. In atropinesterase this residue was found to occupy position 110 in a peptide chain of 272 residues, whereas in trypsin it is located at 177 (length 223), in α-lytic protease at 148 (length 198) and in subtilisin at position 221 (length 275). With regard to general sequence homology, no relationship at all was observed between atropinesterase and the members of these superfamilies. Comparison of atropinesterase with the E.coli enzymes did not reveal any relationship either.

As a further step in the search for possible structural relationship it appeared to be interesting to determine whether any analogy could be observed in the spatial structure of atropinesterase and other serine-enzymes. It was not possible, however, to determine the three-dimensional structure of atropinesterase by X-ray diffraction, owing to the lack of suitable crystals. Only the predicted secondary structure of atropinesterase could be compared with the secondary structures of other serine-enzymes. To predict this structure from the amino acid sequence we used the theoretical method of Lim (78). No resemblance was found between the predicted α-helical regions and β-strands

of atropinesterase and either the predicted or the experimentally determined secondary structures of the "trypsin"-like enzymes and subtilisins. No attempt was made to predict the tertiary fold of atropinesterase, since the prediction theories developed up to date need considerable improvement before any reliable results can be obtained (110).

The absence of any homology between atropinesterase and other known serine-enzymes and the difference in secondary structure indicate the existence of a new "group" of serine-esterases, with -presently- atropinesterase as the only identified member, beside the known superfamilies of subtilisins and prothrombin-related enzymes. The enzymes belonging to this third "superfamily" of serine-enzymes are probably derived from a different ancestral enzyme than the 2 superfamilies known to date. As mentioned before, atropinesterase is a true esterase, whereas all serine-hydrolases known to belong to the other 2 superfamilies, primarily function as a protease in their natural surroundings.

Despite all the functional and structural differences observed, all serine-enzymes seem to possess one equal functional property, i.e. the capacity to hydrolyse ester bonds with the help of one "active serine" residue. This functional similarity indicates that -very likely- atropinesterase has converged from a completely different ancestral sequence to a protein with a functional structure at the active site that is closely related to the corresponding structure of the members of the 2 superfamilies known. This conception is consolidated by the findings of Van der Drift (42), who showed with physico-chemical methods that the active site of atropinesterase shares also other structural elements with the active centers of members of the 2 superfamilies known.

An introduction to the structural analysis of atropinesterase including a description of the strategy followed during sequence determination is given in chapter I. Chapter II briefly reviews the strategy currently used to determine the primary structure of a protein and techniques nowadays available for this purpose. The conventional

methods and the materials, used throughout this study for the determination of the amino acid sequence, for degradation of protein and peptides, for peptide separation and for sequence elucidation are described in chapter III. In chapter IV the methods used to assess the purity of the atropinesterase preparations are given and some characteristics of the protein are described, such as reaction with DFP, isoelectric point, molecular weight, amino acid composition and N- and C-terminal sequence. As a first attempt to establish whether homology exists between atropinesterase and known serine-proteases as trypsin and subtilisin, we aimed at a comparison of the amino acid sequences around the active serine residue. Therefore, our initial efforts were directed at the isolation and sequence determination of peptides containing this active serine, as will be described in chapter V. First a dodecapeptide, obtained after cleavage with CNBr was characterized. In this peptide the active serine was located at position 11, making comparison with the amino acid sequence around the active serine in trypsin and subtilisin unsatisfying. A larger peptide comprising 58 residues, in which the active serine was located at a reasonable distance from either end, was isolated after tryptic digestion of maleylated atropinesterase. The isolation, purification and sequence determination of both peptides are described. In chapter VI the isolation of all 8 peptides obtained after tryptic digestion of maleylated atropinesterase and comprising all residues of the protein is described. Furthermore, the isolation of 6 peptides (comprising 82 residues), obtained after cleavage of the protein with CNBr, and of 2 peptic peptides of the protein is described. The methods to purify these fragments are mentioned and their amino acid compositions are given. In chapter VII all information is presented from which the complete amino acid sequence of the 272 residues-comprising protein could be deduced. The elucidation of the amino acid sequence of the 8 tryptic peptides and the 6 CNBr-peptides is described, together with the isolation and sequence determination of the partial peptides that had to be derived from these fragments to complete their elucidation. Also, the results obtained after digestion of a mixture of large

CN-peptides with various enzymes are given and the results obtained after digestion of the protein with pepsin. Furthermore, the sequence elucidation of some aspecifically produced tryptic peptides is presented and, finally, the assembly of the peptides, which resulted in the identification of all 272 residues along the protein chain. The complete sequence was compared according to the method of Dayhoff (34, 38, 119) with the established sequence of serine-proteases, as is described in chapter VIII. In addition the secondary structure of the protein as predicted by a theoretical method is presented in this chapter, as well as the comparison with the secondary structures, predicted and experimentally determined, of some serine-proteases from the trypsin and subtilisin families.

## CHAPTER II ADVANCES IN PROTEIN SEQUENCING

Strategies and techniques commonly used in protein-sequencing have been published in excellent general reviews of methodology (89, 90, 65, 73), in specific reviews of technology (91, 92, 99) and in shorter discussions of strategy (133, 21), while advances in strategy and methodology have been reviewed by Walsh et al. (132). A general introduction into the strategy and tactics of protein sequence determination is therefore omitted in this thesis. Only some recent developments will be described which have led to the increase in rate of accumulation of amino acid sequence data. In 1976 Edman (48) estimated that the elucidated sequences comprised about 80,000 amino acid residues. In 1979 this number had doubled and about 160,000 residues had been identified in about 1100 completely sequenced polypeptide chains. This rapid growth was mainly caused by the introduction of automated sequential degradation technology as was realized in the Beckman sequencer. In April 1982 the number of residues located in known sequences had almost doubled again and comprised 289,574 residues in 1809 proteins (M.O. Dayhoff et al., personal communication from the current data base of the Atlas of Protein Sequence and Structure (37)). This further increase in sequence data was mainly due to the introduction of new techniques for peptide separation and improvement of the Edman degradation procedures, both automatically and manually, while advances in DNA-sequencing have also contributed to the accelerated accumulation of these amino acid sequence data.

From the reviews mentioned above it is clear that the methods for protein sequencing are well worked out. However, it is difficult to make the proper choice out of the many possibilities with respect to fragmentation of the protein, peptide separation, detection of peptides and amino acid sequencing procedures. No general strategy can be given for any given protein. For instance, whether application of distinct fragmentation methods will be successful depends on the availability of the starting material, the size of the protein, the occurrence of specific amino acids in the protein and on the distri-

bution of these specific residues along the polypeptide chain. Usually, only during the sequencing -when more information on the protein studied accumulates- the insight is gained into what would have been the best choice. Furthermore, the availability of technical equipment in a laboratory plays a role in the strategy to be followed, but in a laboratory that possesses all the modern equipment necessary for efficient, sensitive and reliable sequence determinations, nowadays one would commonly prefer two or more restrictive specific agents for fragmentation of the protein, high-performance liquid chromatography (HPLC) for peptide separation, sequence determination by automated sequencer procedures and identification of amino acids as phenylthiohydantoin-derivatives by HPLC procedures. Nevertheless, even under these ideal conditions problems may be met that have to be solved by the application of less sophisticated techniques, as many proteins contain sequences that refuse to comply with the intended course of action, e.g. because insoluble fragments are formed or fragments that adhere persistently to column material, or because particular peptide bonds show an abnormal (in)stability.

## II.1 FRAGMENTATION PROCEDURES

Specific restrictive cleavage techniques for degradation of proteins into appropriate fragments are used nowadays as a first step in protein sequence determination to produce a limited number of peptides. They have replaced classical enzymatic and chemical cleavage procedures since, firstly, it is more convenient -in principle- to separate a small number of large peptides than a large number of smaller ones and, secondly, these techniques are more productive in concert with automated Edman degradations. The methods currently used comprise cleavage procedures with chemical or enzymatic agents which split specific peptide bonds, (aspartyl-prolyl; asparaginyl-glycyl) or disrupt the polypeptide chain at specific residues (at methionyl(Met), at tyrosyl(Tyr) and tryptophanyl(Trp), at lysyl(Lys), at arginyl(Arg) or at acidic residues) or which split the native enzyme at specific sites. The classical cleavage techniques, such as cleavage with

trypsin, pepsin, thermolysin and partial acid cleavage, are still applied for the degradation of large peptides.
Most of the specific restrictive fragmentation agents have been known already for a long time, but their application has awaited the introduction of sequencing techniques that require large fragments.

## II.2 SEPARATION METHODS

In contrast to the fragmentation procedures, where the developments were less in the techniques than in the preferences, much improvement has been achieved in the methods used for separation of peptides by the introduction of HPLC. Yet, conventional methods such as paperchromatography, paperelectrophoresis, gelfiltration on Sephadex columns, chromatography on cation- or anion-exchange columns are still used in sequence studies, owing to the fact that application of HPLC is far from routine. The growing interest in HPLC is due to the availability of efficient column materials consisting of silica-based microporous particles (5-10μm) with bonded nonpolar moieties. These columns are used in "reversed-phase" chromatography in which the eluent is more polar than the stationary phase. Reversed phase HPLC has advantages of speed, sensitivity and high resolving power when compared with other separation techniques. Further development of this technique and improvement of resins makes this technique the method of choice in current and future sequence studies.
As for the detection of peptides in column effluents and on paper no new reagents have been introduced in the last five years.

## II.3 PROCEDURES FOR DETERMINATION OF THE AMINO ACID SEQUENCE

For determination of the amino acid sequence of proteins and peptides, the Edman degradation is still the method of choice.
The classical sequence procedures applied for amino acid sequence determination of small fragments, such as the manual Edman degradation and the more sensitive dansyl-Edman degradation, were extended some 10 years ago with the automated Edman degradation procedures with the

Beckman "spinning cup" sequencer and the solid phase sequencer. The introduction of these techniques, which reduced the time required for sequence elucidation, did not eliminate the need of manual procedures on small fragments, since these new techniques were only applicable on a few hundred nanomoles of protein or peptide and on fragments of at least 100 and about 30 residues, respectively. Beside this aspect, general application was also limited by the high costs of installment, continuous operation and maintenance of the equipment. The recent adaptation of the Beckman sequencer to the degradation of small amounts (nmol quantities) of fragments of any size, together with improved detection methods meant a striking improvement in one of the last steps of protein sequence determination. Manual procedures, which are still being used for small peptides, have been supplemented with a more sensitive method, in which 4-N,N-dimethylaminoazobenzene 4'-isothiocyanate is used as the coupling reagent in stead of phenyl isothiocyanate (PITC). All these developments have led to faster and more sensitive sequence procedures.

## II.4 METHODS FOR IDENTIFICATION OF AMINO ACID RESIDUES

In sequencing procedures based on the manual degradation method, amino acids were identified either as dansyl-derivatives after thin layer chromatography (TLC) and detection of the fluorescent products, or as phenylthiohydantoin (PTH-)amino acids after TLC and detection of the residues with more or less sensitive staining reagents. PTH-amino acids produced by automated methods were detected in the same way. Since the introduction of reversed-phase HPLC systems, this technique is the method of choice for identification of PTH-amino acids, produced by either manual or automated degradation, because of its high speed and low detection limit (pmol quantities). In the sequence study described here conventional but gradually updated techniques have been used for the determination of the majority of the amino acid sequences, while at the final stage use was made of more sofisticated techniques with modern equipment in other laboratories.

## II.5 FUTURE ASPECTS FOR PROTEIN SEQUENCE WORK

Improvement of separation techniques and the adaptation of automated degradation with the Beckman sequencer for small peptides has considerably speeded up the elucidation of protein sequences. However, the time needed for elucidation still markedly increases when larger proteins are involved or when only small amounts of starting material are available.

Elucidation of the primary structure of proteins will continue to provide a chemical base for understanding proteins in their role in biology, but in coming years much of the information on the polypeptide backbone of proteins will be provided by the even more rapid DNA-sequencing technology in cases that identification and cloning of the encoding gene is possible. The DNA sequence, however, does not give all information on the functional protein, as it does not reveal the sites of disulfide bridges, or the possible removal of peptide segments during "processing" reactions.

Cooperation between protein sequence chemists, DNA-sequencing groups and X-ray crystallographers appears to be of great importance to get an answer to the question what is the magic role played by the 20 amino acids in living organism, for which -furthermore- great benefit is to be expected from the growing possibilities to introduce at will changes in the polypeptide chain, either by protein chemistry or -more importantly- by genetic engeneering.

# CHAPTER III METHODS AND MATERIALS

## INTRODUCTION

In this chapter materials and methods of a more general nature, used in this study, are described. More detailed information concerning the particular experiments and a description of specific methods will be given in the chapter in question (IV through VII) under the heading "Experimental details".

## III.1 METHODS

### III.1.1 METHODS USED TO ISOLATE AND CHARACTERIZE ATROPINESTERASE

Molar amounts or concentrations given refer to atropinesterase monomers (=active sites).

III.1.1.1 Assay of atropinesterase activity was performed as described by Rörsch et al.(108). The enzymatic activity of the protein sample was determined by acidimetric titration of tropic acid, released during the hydrolysis of the substrate (—)-atropine, in a pH-stat. The Radiometer (Copenhagen) Titrator TTT 1b and Titrigraph SBR 2c were used, which automatically provide for the addition and registration of the amount of base needed per unit of time to maintain a constant pH. Activity determinations were carried out at pH 7.0; the enzyme (1 to 10 µl) was added to 10 or 25 ml of 0.4 mM (—)-atropine sulfate in 0.1 M KCl and 0.02% saponine in a vessel thermostatted at 25°C, and the consumption of 0.02 M NaOH was registered for at least 5 minutes. The activity of the enzyme solution is expressed as units (U) per ml. One unit of enzyme is defined as the amount that will cause the hydrolysis of 1 µmole of (—)-atropine per minute under the above conditions. The specific activity (SA), a measure for the purity of the preparation, is defined as the number of units per A280. (The relation A280 = ca. 0.5 mg protein per ml was determined by amino acid analysis). The SA of (almost) pure samples varied between 500 and 600 U per A280.
The method used to determine the activity of atropinesterase is simple, reliable and deviation of zero order kinetics only occurs after more than 75% of the substrate has been consumed.

III.1.1.2 Isolation and purification of atropinesterase. Pseudomonas putida bacteria (strain L, PMBL 1) were grown in a semi-continuous culture in a cumulative volume of ca.600 l., resulting in several kg (wet weight) of cells. Atropinesterase was isolated from these bacteria and purified at our laboratory, according to the method of Rörsch et al. (108) as modified by Oosterbaan (to be published). Fractions with a SA between 500 and 600 U per A280 were used.

III.1.1.3 Inhibition of atropinesterase by [$^{32}$P]-diisopropyl phosphorofluoridate (DFP) was performed essentially in the same way as the inhibition with Soman (1,2,2-trimethylpropyl methylphosphonofluoridate), described by Rörsch et al. (108). Atropinesterase, activity 1000-2000 U/ml (SA 500-600), was incubated in 0.03 M Tris.HCl, pH 8.2, with a ca. 5 fold molar excess of [$^{32}$P]-DFP (over DFP-binding groups in the enzyme) for 4-24 h at room-temperature. The amount of $^{32}$P bound to the protein was determined in samples taken during inhibition and after complete inactivation. To this end, the samples were applied to a paperchromatogram (Whatman no.1) which was developed (ascending) with a mixture of 90 parts of n-butanol, 11 parts of glacial acetic acid and 30 parts of water (BAW). After drying, the radioactivity which had remained at the application area (corresponding to the protein-bound $^{32}$P) was counted. After (almost) complete inhibition the solution was dialyzed against water at 4°C for 48 h to remove excess [$^{32}$P]-DFP and other $^{32}$P-products, mainly diisopropyl phosphate (DIP), the hydrolysis product of DFP. The effectiveness of the dialysis and the amount of $^{32}$P bound to the protein after dialysis were also determined by paperchromatography. As complete removal of DIP from protein solutions is very difficult (11), the final preparation still contained some unbound $^{32}$P, varying between 5% and 50% of the total $^{32}$P-content.

Determination of $^{32}$P-radioactivity:

1) To determine the $^{32}$P-radioactivity on paperchromatograms we used a "scanning counter", i.e. an end-window Geiger Müller counting tube (General Electric Co. England, type GM4, connected with a Philips radiation registration apparatus PW 4035) vertically mounted, and adjustable in 3 directions, above a plate on which the chromatogram was placed (efficiency of $^{32}$P-counting ca. 15%).

2) The $^{32}$P-radioactivity in fluids was determined by counting aliquots in a liquid scintillation counter (Nuclear Chicago Mark I) according to the method of Chérenkov.

III.1.2 SEPARATION METHODS ON POLYACRYLAMIDE GELS

Isoelectric focusing (IEF) and polyacrylamide gelelectrophoresis (PAGE) were used to control the purity of atropinesterase. Furthermore these techniques were used to determine the isoelectric point (IEF) or to determine the molecular weight (PAGE).

III.1.2.1 Polyacrylamide gelelectrophoresis (PAGE) without sodium dodecyl sulfate (SDS) was performed according to Maurer (83) and with SDS by a combination of the method of Dunker and Ruekert (44) and Weber and Osborn (134).

Preparation of gels.

a) 7% Polyacrylamide gels without SDS, prepared according to Maurer, contained 7% acrylamide, 0.03% N,N,N',N'-tetramethylethylenediamine (TEMED), 0.18% methylene-bisacrylamide (Bis), 0.14% ammonium persulfate (all w/v) and 0.38 M Tris.HCl, pH 8.9. A Tris-glycine buffer, pH 8.3, containing 1.2 g Tris and 2.88 g glycine per 1, was used as electrophoresis buffer.

b) 10% Polyacrylamide gels with 0.1% SDS were made as described by Dunker et al. The 10% acrylamide solution was mixed with 5% by volume of 1% ammonium persulfate and contained 0.1% SDS, 0.05 M sodium phosphate, pH 7.2, and 0.10% (v/v) TEMED. The acrylamide to Bis ratio was 29:1 (w/w). The electrophoresis buffer contained 0.1 M sodium phosphate, pH 7.2, and 0.1% SDS. The gels were cast in plastic tubes of about 7 cm length with an inner diameter of 8 mm.

Preparation of samples for electrophoresis.

a) for gels without SDS: to atropinesterase samples and marker proteins, dissolved in ca. 50 µl 0.03 M Tris.HCl, pH 8.15, containing 0,02 M KCl, and water, respectively, 10 µl glycerol was added to facilitate application and to prevent mixing with buffer. Protein concentrations were ca. 2 mg/ml.

b) for gels with SDS: atropinesterase and marker proteins were dissolved at a concentration of 1-2 mg/ml in 4 M urea and about 1% SDS, while in some experiments 1% 2-mercaptoethanol (v/v) was added. The solutions were incubated at 45°C for 45 min.

ad a+b) ca.50 µl of each protein solution was applied to one gel. To all samples 5 µl 0.05% Bromophenol Blue in water was added as a reference compound.

Electrophoresis was performed at ca. 5mA and 20 V per gel for about 4 h.

Staining and destaining was performed as described by Weber and Osborn (134). Gels were immersed in a solution of 1.25 g Coomassie Brilliant Blue R250 in 454 ml 50% methanol and 46 ml acetic acid for 2-16 h, at roomtemperature, then rinsed with water and destained in a solution of 75 ml acetic acid, 50 ml methanol and 875 ml water. Destaining was completed either by changing the solution 2-3 times over a 24 h period or by transverse electrophoresis for 1-1½ h.

III.1.2.2 Isoelectric focusing (IEF) was performed on polyacrylamide sheets, purchased from LKB, with Ampholine carrier ampholytes in the range covering pH 3.5 to 9.5, as described by LKB in a LKB Multiphor apparatus. Samples (15-20 µl) of atropinesterase (ca.1500 U/ml) were applied into small incisions made in the middle of the sheet. Staining with Coomassie Brilliant Blue R250 and destaining with ethanol-acetic acid-water (50:16:134, by volume) were performed according to the LKB manufacturer's manual. The pH gradient in the gel was measured with a micro glass electrode.

## III.1.3 MODIFICATION OF PROTEIN AND/OR PEPTIDES

### III.1.3.1a Modification of atropinesterase with 1,4-[$^{14}$C]-maleic anhydride and denaturation of the maleylated protein.

Atropinesterase (0.02-3.0 µmol; in some cases [$^{32}$P]-DFP-labeled) was dialyzed against 2 l 0.01 M sodium borate buffer, pH 9.0 for 24 h at 4°C, then concentrated by lyophilization to 2.1 mg protein per ml (70 nmol/ml) in 0.02 M sodium borate, pH 9.0. Maleylation was essentially as described by Butler et al. (23). Preparations of 50 µCi 1,4-[$^{14}$C]-maleic anhydride (specific radioactivities of 14.3, 28.8 and 49 mCi per mmol were used) were dissolved in a dry benzene solution of car-

rier maleic anhydride, which had been purified by sublimation under reduced pressure, to give solutions of 0.3 M of 1,4-[$^{14}$C]-maleic anhydride with specific radioactivities ranging from 11-40 μCi/mmol. (The radioactivity found in these solutions was always much less than the $^{14}$C-content of the vial indicated by the supplier; it varied from 14 to 60% of that amount, presumably because of partial hydrolysis of the [$^{14}$C]-maleic anhydride delivered). The solution with the lowest specific radioactivity was used to modify the protein from which the active serine-containing tryptic peptide, Tm-5, was isolated. Solutions with higher specific radioactivities were used for preparations from which other tryptic peptides were isolated as well.
To modify the protein, a volume of the 0.3 M solution of 1,4-[$^{14}$C]-maleic anhydride, amounting to a 25 molar excess over the total amount of free amino groups present in the protein, was added to the solution of atropinesterase, in 2 or 3 portions over a period of 15 min, at 2°C. The pH of the reaction mixture was maintained between 8.5 and 9.0 by the addition of 4 N NaOH. After 30 min of stirring at roomtemperature, the benzene was blown off with nitrogen in a ventilated hood until a clear solution was obtained. The ([$^{32}$P]-DFP-labeled) maleylated atropinesterase was desalted by passing the solution down a column of Sephadex G-50 superfine grade (sf) in 0.01 M Tris.HCl, pH 8.0. It has to be emphasized that the sample was never concentrated by lyophilization prior to application to the column in order to prevent undesired fragmentation of the maleylated protein, because in previous experiments lyophilization of these solutions seemed to cause heterogeneity in the protein material (definite proof that this was indeed the cause was not obtained, however). After chromatography, the maleylated protein was recovered in a yield of 92% according to the A280. Fractions containing the modified protein were pooled and concentrated by lyophilization to a concentration of about 115 nmol/ml. The maleylated protein was then denatured by heating in boiling water for 5 min and after cooling at roomtemperature the solution was divided in several portions, which were stored at -20°C until they were used for -mainly- tryptic digestion. This way of storage avoided repeated freezing and thawing of the same protein solution, which appeared to cause breaks in the peptide chain.

III.1.3.1b Removal of 1,4-[$^{14}$C]-maleylgroups from α- and ε-amino groups in peptides. Peptide material was concentrated or dried by lyophilization; subsequently acetic acid was added to give a concentration of 150-200 nmol peptide per ml 33% acetic acid. Demaleylation was performed by incubation at 40°C for 48-68 h. (pH about 2).

III.1.3.2 Carboxymethylation of reduced atropinesterase and of a (reduced) tryptic peptide Tm-5.

III.1.3.2a Atropinesterase (265 nmol; 8 mg) was dissolved in 1.3 ml 0.02 M Tris.HCl, pH 8.5, containing 8 M urea. Reduction was performed with dithioerythritol (DTE) as described by Cleland (29): 5.2 mg DTE in 0.65 ml 0.02 Tris.HCl, pH 8.5, 8 M ureum and 3 mM EDTA was added and allowed to react for 2 h at 37°C under nitrogen. Then

carboxymethylation was performed, by the method of Crestfield (33) with slight modification. To 1.9 ml of the reaction mixture an equal volume of 0.1 M iodoacetic acid was added under a stream of nitrogen, in the dark. Simultaneously the pH of the reaction mixture was maintained at pH 8.5 by addition of 0.5 N NaOH. After 1 h under nitrogen at roomtemperature in the dark the reaction was stopped by the addition of 75 µl 2-mercaptoethanol. The solution was desalted on a Sephadex G-50 sf column (54 x 2 cm) equilibrated with 0.01 M Tris.HCl, pH 8.0. The fractions with carboxymethylated protein were pooled and concentrated by lyophilization.

III.1.3.2b Demaleylated peptide Tm-5 (1.6 µmol; 10 mg) was reduced and carboxymethylated in an analogous way as described for the protein, at an initial peptide concentration of 340 nmol/ml. For carboxymethylation 0.1 M 2-[$^{14}$C]-iodoacetic acid (specific radioactivity 151 µCi/mmol), pH 8.5, was used. The reaction was stopped by addition of 0.45 ml 2-mercaptoethanol. After dilution with 14 ml glacial acetic acid, the reaction mixture was desalted on a Sephadex G-50 sf column equilibrated with 50% acetic acid.
This procedure was applied to 94.8% of the peptide; the remaining 5.2% was treated identically but for the reduction with DTE which was omitted.
The specific radioactivity of the 0.1 M 2-[$^{14}$C]-iodoacetic acid solution was determined via conversion into 2-[$^{14}$C]-glycine. This solution had been obtained by dissolving an amount of material supposed to contain 2.78 µmol 2-[$^{14}$C]-iodoacetic acid with a specific radioactivity of 54 mCi/mmol in 7.75 ml 0.1 M iodoacetic acid, titrated to pH 8.5 with 4 N NaOH; an 0.2 ml aliquot was added to 0.2 ml ammonia (14.7N) and left at roomtemperature for 2 days. The reaction product 2-[$^{14}$C]-glycine, which should have the same specific radioactivity as the iodoacetic acid, was -divided over 3 spots- purified by high-voltage paperelectrophoresis at pH 3.6 for 90 min. After staining of guidestrips with ninhydrin, spots corresponding to standard glycine were eluted. After determination of the radioactivity in a certain volume of the eluate by liquid scintillation counting and after determination of the amount of glycine in an equal volume by amino acid analysis, the specific radioactivity was calculated to be 151 µCi/mmol, 78% of the theoretical value.

III.1.4.1 AMINO ACID ANALYSIS
All amino acids are indicated by the generally used three-letter code given in the list of abbreviations on page 8.
Protein or peptide samples up to 20 nmol were hydrolyzed in 6 N HCl in Pyrex tubes. Prior to use, these tubes had been filled with 0.4 ml 6 N HCl, sealed, heated at 110°C for 18 h, opened, rinsed and dried. Norleucine was added to the samples as an internal standard. Water was removed by lyophilization, 0.4 ml 6 N HCl was added, tubes were flushed 3 times with nitrogen, cooled in a bath of solid $CO_2$ in ethanol, evacuated and then sealed (1); hydrolysis was for 18-24 h at 110°C, unless otherwise stated. In the various tables no corrections were made for hydrolytic losses of Thr and Ser unless otherwise mentioned;

the same applies to possible incomplete recovery of residues present in sequences that need prolonged hydrolysis for complete scission. When Trp had to be determined, 2% (by volume) thioglycolic acid (TGA) was added to the 6 N HCl following the method of Matsubara (81), which gave a recovery of about 90%. Addition of TGA also prevented the oxidation of Met, Tyr and carboxymethylcysteine (CMCys). When CMCys had to be determined, 2% phenol (by volume) was added to the 6 N HCl and extra care was taken to exclude oxygen. After hydrolysis, HCl was removed by evaporation at 30°-50°C in vacuo. When homoserine (Hse) was present the dried hydrolysate was incubated with 0.1 ml pyridine-acetic acid-water (25:1:225, by volume) pH 6.5 buffer at 110°C for 1 h to convert all Hse-lactone to Hse (5).
After lyophilization, samples were stored in the dry state or analyzed immediately. Initially, amino acid analyses were performed on a Beckman model 120 amino acid analyzer according to the two-column system of Spackman et al. (122) with the procedures described by Spackman (123). Most of the analyses, however, were carried out on an LKB 3201 amino acid analyzer with a one-column system (diameter 0.9 cm). Routinely, buffers of pH 3.25, 4.25 and 6.25 were used (LKB manual), but a first eluting buffer of pH 3.10 was used when Hse had to be separated from Glu. The amino acid analysis system did not resolve Gln and Asn from Thr or Ser but the presence of amides could generally be concluded from the unusually high absorption at 440 nm, relative to 570 nm, of the coloured product formed by the amide with ninhydrin (cf. Glu and Lys; 7). (It has to be noted that elution positions of Asn and Gln slightly changed with minor -uncontrolled- variations during operation of the amino acid analyzer). Standarized amino acid reference samples, prepared by Organon*, were used throughout this study.

III.1.4.2 AMINO ACID COMPOSITION

The amino acid composition of the protein was determined from the amino acid analysis and the estimated molecular weight of the monomer, i.e. 30,000.
The amino acid composition of peptides was estimated from the amino acid analysis and (i) the amount per residue released by exopeptidases (see III.1.5.1a and III.1.5.2), and/or (ii) the molecular weight as roughly estimated after chromatography on Sephadex G-50. In many cases determination of the composition was facilitated because it could be assumed that the peptide would contain only one residue of (a) certain amino acid(s) for instance as the consequence of the particular degradation procedure used (e.g. C-terminal Arg) (III.6.1). As a rule, the composition of smaller peptides could be determined entirely, simply by looking at the ratio. The amino acid compositions of peptides given in the tables are derived by taking the nearest integer of the ratio number of each amino acid. Ratio numbers were found by dividing the quantity of each amino acid as found by amino acid analysis, by a constant factor, which was derived from the sum of the quantity of each amino acid divided by the total number of residues as identified by one of the above methods or in a later stage by sequence analysis.

*Thanks are due to Ir. B.C.Goverde, Organon, Oss and SON for this gift.

III.1.5 ENDGROUP DETERMINATION

N-terminal analysis of protein and peptides was performed enzymatically with leucinaminopeptidase (LAP), whereas dansylchloride was used to identify the N-terminal amino acid in a chemical way. The exopeptidases carboxypeptidase A and/or B (CPA and/or CPB) were used to identify one or more residues located at the C-terminus of protein or peptide. Amino acids released by LAP, CPA and/or CPB were identified by amino acid analysis.

III.1.5.1a N-terminal end group analysis with LAP. Peptides (5-20 nmol) were incubated in 50 µl 0.05 M Tris.HCl, 2.5 mM $MgCl_2$, pH 8.5, with 10 µl LAP (1-4 mg/µmol of peptide) at 37°C for various periods (77). The LAP solution (2,4 mg/ml; 18.2 U/mg) in 3.5 M $(NH_4)_2SO_4$ pH 7.7, (containing 10 mM $MgCl_2$) was preincubated for 1 h at 37°C. Reaction was stopped by addition of 20 µl pH 2.2 sodium citrate buffer, containing norleucine (7 nmol) as an internal standard. Samples were analyzed after dilution with 0.4 ml pH 2.2 sodium citrate buffer (77). The results were corrected for possible losses on the basis of the recovery values of the norleucine added. A LAP-mixture without peptide was run through the whole procedure as a blank to check whether a background of free amino acids was present.

III.1.5.1b N-terminal end group determination with dansylchloride. The N-terminal residue of peptides was identified essentially by the method of Hartley (60) as described under III 2.10.1b. The same procedure was applied to identify the N-terminal residue of the protein, but slightly modified by the addition of 1% SDS (58). Details are given under IV.6.1.

III.1.5.2 C-terminal end group analysis by digestion with CPA and/or CPB followed the method of Ambler (4). The protein or peptide (5-20 nmol), was dissolved in 70 µl 0.2 M N-ethylmorpholinium acetate, pH 8.5, and digested with 20 µl CPA and/or CPB solution (1 mg/ml; 1-4 mg/µmol of peptide or protein) at 37°C for different periods. Stopping of the reaction, further treatment and control experiments were as described sub III.1.5.1a.
Carboxypeptidase solutions were prepared by washing 0.125 mg CPA and/or CPB suspension with 0.5 ml water; after centrifugation for 5 min at 700 x g, the supernatant was discarded; the crystals were suspended in 10 µl 1% (w/v) $NaHCO_3$ and dissolved with 20 µl 0.1 N NaOH. After adjustment of the pH to 8-9 with 20 µl 0.1 N HCl, 70 µl 0.2 M N-ethylmorpholinium acetate, pH 8.5, was added to give a final enzyme concentration of 1 mg/ml. Solutions were freshly prepared before use. Sometimes, different CPA preparations differed markedly in their action upon acidic residues. Some preparations did not release Asp or Glu, not even after extensive hydrolysis, which agrees with the claimed specificity of pure preparations, whereas other preparations liberated these residues very easily. This aspecific action of some samples was used gratefully to localize acidic amino acids in peptides.

### III.1.6 CHEMICAL AND ENZYMATIC METHODS USED FOR DEGRADATION OF MALEYLATED ATROPINESTERASE

NOMENCLATURE:
Peptides derived from digestion of maleylated atropinesterase (Am) with trypsin were designated as Tm-, with clostripain as Cm-, fragments resulting from CNBr-cleavage of atropinesterase (AE) as CN- and fragments resulting from peptic cleavage of the protein as P-. An unresolved mixture of large CN-peptides is designated as "CN-peptides", whereas a complex mixture of tryptic peptides from maleylated atropinesterase produced by uncontrolled rupture of peptide bonds, is referred to as [Tm-]. Peptides derived from tryptic cleavage of Tm- and CN-peptides are designated as T-, from peptic digestion as P-, from digestion with staphylococcal protease as SP-, from thermolytic digestion as Th-, from chymotryptic digestion as Ch- and from NBS-cleavage as NBS-. Peptides derived from CNBr-cleavage of Tm-peptides are designated as CNBr-, while peptides derived from tryptic, peptic and staphylococcal protease digestion of impure "CN-"fragments are designated as t-, p- and sp-, respectively.

III.1.6.1 Cleavage of atropinesterase with cyanogen bromide (CNBr). Methionylbonds were cleaved with CNBr according to Gross (59). The lyophilized protein (0.05-2.5 µmol; in some experiments labeled with [$^{32}$P]-DFP) was dissolved in 70% (v/v) formic acid containing a 180-fold molar excess of CNBr relative to Met to give a final protein concentration of 11 mg/ml (0.36 mM) and a CNBr concentration of 60 mg/ml (0.58 M; table III.1). After 6 h incubation at 25°C under nitrogen in the dark, the material was fractionated on a column (160 x 0.9 cm or 93 x 2.5 cm) of Sephadex G-50 sf. After 6 h incubation the reaction appeared to be complete: no Met was found in the hydrolysate of the resulting peptide mixture, indicating that all Met residues were converted into Hse or Hse-lactone. After reaction times longer than 6 h no change was observed in the elution profile on Sephadex G-50 sf obtained by monitoring the effluent for the presence of peptides according to the procedures described in III.1.9.1.

III.1.6.2 Digestion of maleylated protein with trypsin. In an early stage of this study 20-100 nmol [$^{32}$P]-DFP-labeled maleylated atropinesterase was digested at a protein concentration of 3 mg/ml in 0.03 M Tris.HCl, pH 8.0, with 3% (by weight of substrate) bovine L-(tosylamido-2-phenyl)ethyl chloromethyl ketone (TPCK-)treated trypsin (Merck) for 6 h at 37°C (table III.1). TPCK-trypsin was added in two equal portions at zero time and after 3½ h of incubation at 37°C. The peptide mixture was immediately applied to a column (150 x 0.9 cm) of Sephadex G-50 sf. In subsequent experiments the maleylated protein (0.05-3.0 µmol) was digested at a protein concentration of 3.0-3.6 mg per ml (100-120 nmol/ml) in 0.05 M Tris.HCl, pH 8.0, with 0.5% (by weight of substrate) porcine trypsin (Miles; 1 mg/ml 0.001 N HCl) for 10 min at 37°C. Reaction was stopped by heating the solution for 3 min at 100°C. The peptide mixture was applied to a column (136 x 2.3 cm) of Sephadex G-50 sf. When the volume to be fractionated exceeded

17 ml, the digest was divided in two portions, and the second portion was applied to the column after storage overnight at -20°C.

III.1.6.3 Digestion of maleylated atropinesterase with clostripain (86) was performed with 0.84 µmol maleylated atropinesterase at a concentration of 115 nmol/ml. The peptide mixture was fractionated on a column of Sephadex G-50 sf. Further information is given in VI-I.1.1.

III.1.6.4 Digestion of atropinesterase with pepsin was performed with 0.65 µmol atropinesterase in 5% formic acid at a concentration of 200 nmol/ml for 4 h at 25°C (table III.1).

III.1.7 SECONDARY CLEAVAGE OF (DEMALEYLATED) TRYPTIC, CLOSTRIPAIN, AND CYANOGEN BROMIDE PEPTIDES OR OF FRAGMENTS OF THESE PEPTIDES BY ENZYMATIC OR CHEMICAL PROCEDURES

Details about the amount of peptide cleaved, the peptide concentration, the amount of enzyme or chemical added and the incubation times are given in table III.1.

III.1.7.1 Tryptic digestion of CN-1 and CN-8, of "CN-"peptides, and of the demaleylated peptides Tm-5, Tm-6, Cm-6a, Tm-7, Tm-8 and Tm-1b was performed in 0.02 M Tris.HCl, pH 8.0 or 8.5, at a peptide concentration of 150-1150 nmol/ml, for periods varying from ½ to 16 h, at 37°C, with ca. 0.1-0.2 mg trypsin per µmol peptide. Trypsin was added at zero time and sometimes again after 1½ or 2 h of incubation. Reaction was stopped by lyophilization or acidification.

III.1.7.2 Peptic digestion of demaleylated Tm-1, Tm-2, Tm-6 and Tm-8 was carried out in 5% formic acid at a peptide concentration of about 200 nmol/ml, for 1-1½ h at 25°C. Usually 0.25 mg pepsin per µmol peptide were used. The reaction was stopped by lyophilization or acidification. Tm-5 was taken up in 5% formic acid, resulting in a slightly turbid solution with a total peptide concentration of 225 nmol/ml. Digestion of this peptide was performed for 1½ h at 25°C with 0.6 mg pepsin per µmol peptide. After digestion the solution could not be clarified with a few drops of 96% formic acid or by centrifugation for 15 min at 40,000 rpm (300,000 x g). The insoluble material accumulated on top of the clear solution but mixed with the solution on touching the tube. The unclear solution was fractionated without further treatment.

III.1.7.3 Digestion with staphylococcal protease (S.protease) of demaleylated Tm-2, Tm-6, Tm-6 T-2, Tm-7 and [Tm-8b], and of CN-1, CN-8 and "CN-"peptides was carried out in 0.05 M phosphate buffer, pH 7.7, or in 0.05 M NH $HCO_3$, pH 7.8, (66) at a peptide concentration of 275-600 nmol/ml, for 16 or 21 h at 37°C, with 0.1-1.5 mg S.protease per µmol peptide. S.protease was added as a "solution" of 2.5 mg/ml in 0.05 M phosphate buffer, pH 7.7. The reaction was stopped by lyophilization.

Table III.1 Summary of the conditions used in the enzymatic and chemical fragmentation of atropinesterase and of the various peptides.

| Peptide | | | conditions of incubation | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| code name | amount (nmol) | concentration (nmol/ml) | duration (h) | temp. (°C) | amount of agent (µg) | (mg/µmol) | (%)* | 2nd add. after (h) | termination by |
| TRYPSIN | | | | | | | | | |
| DFP-Am | 20- 100 | 100 | 6 | 37 | | | 3.0 | 3½ | heating |
| Am | 50-3000 | 100-120 | 1/6 | " | | | 0.5 | - | " |
| [Tm-1b] | 50 | 500 | 16 | " | 2x 5 | 0.20 | 11 | 1½ | lyophil. |
| Tm-5 | 230 | 250 | 16 | " | 2x 25 | 0.22 | 3.6 | 2 | acidif. |
| Cm-6a | 550 | 250 | 4 | " | 2x 45 | 0.16 | 2.3 | 1½ | lyophil. |
| Tm-6 | 400 | 150 | 1 | " | 40 | 0.10 | 1.4 | - | " |
| Tm-7 | 400 | 400 | ½ | " | 50 | 0.13 | 3.1 | - | " |
| Tm-8 | 225 | 640 | 1 | " | 18 | 0.08 | 2.0 | - | acidif. |
| CN-1 | 350 | 1150 | 5 | " | 2x 5 | 0.09 | 3.0 | 1½ | " |
| CN-8 | 150 | 1150 | 5 | " | 2x 5 | 0.08 | 3.2 | 1½ | " |
| "CN-"pept. | (10)** | (10)** | 6 | " | 2x150 | | 3.0 | 2½ | lyophil. |
| PEPSIN | | | | | | | | | |
| AE | 670 | 200 | 4 | 25 | 200 | 0.30 | 1.0 | | freezing |
| Tm-1 | 200 | 200 | 1 | " | 20 | 0.10 | 3.7 | | lyophil. |
| Tm-2 | 340 | 200 | 1 | " | 85 | 0.25 | 10.0 | | acidif. |
| Tm-5 | 200 | 225 | 1½ | " | 130 | 0.60 | 10.0 | | " |
| Tm-6 | 550 | 225 | 1½ | " | 75 | 0.25 | 3.5 | | formic acid |
| Tm-8 | 133 | 200 | 1 | " | 33 | 0.25 | 6.0 | | acidif. |
| "CN-"pept. | (40)** | (27)** | 2 | " | 520 | | 1.5 | | |
| THERMOLYSIN | | | | | | | | | |
| Tm-7 | 150 | 750 | 3 | 37 | 15 | 0.10 | 3 | | lyophil. |
| [Tm-1b] | 110 | 240 | 3 | " | 17 | 0.15 | 10 | | " |
| CN-1 | 150 | 1000 | 6 | " | 10 | 0.07 | 2 | | " |
| CN-3 | 300 | 1000 | 2 | " | 10 | 0.03 | 2 | | " |
| CN-8 | 150 | 1000 | 6 | " | 5 | 0.03 | 1.2 | | " |
| S.PROTEASE | | | | | | | | | |
| Tm-2 | 205 | 600 | 16 | " | 75 | 0.4 | 15 | | " |
| Tm-6 | 305 | 190 | 16 | " | 138 | 0.5 | 6 | | " |
| Tm-6 T-2 | 32 | 450 | 16 | " | 5 | 1.5 | 12 | | " |
| Tm-7 | 133 | 275 | 16 | " | 88 | 0.7 | 17 | | " |
| [Tm-8b] | 200 | 400 | 16 | " | 38 | 0.2 | 8 | | |
| CN-1 | 500 | 500 | 16 | " | 40 | 0.1 | 2 | | " |
| CN-8 | 350 | 350 | 16 | " | 80 | 0.2 | 9 | | " |
| "CN-"pept. | (11)** | (10)** | 21 | " | 500 | | 5 | | " |
| CHYMOTRYPSIN | | | | | | | | | |
| Tm-7 SP-1 | 240 | 240 | 3 | " | 30 | 0.12 | 4 | | " |
| Tm-3 | 60 | 240 | 2 | " | 2 | 0.25 | 1.6 | | |
| NBS | | | | | | | | | |
| Tm-7 | 135 | 270 | 1 | 25 | | | 2 | | " |
| CNBr *** | | | | | | | | | |
| AE | 50-2500 | 360 | 6 | " | | | | | |
| Tm-5 | 100 | 80 | 17 | " | | | | | gelfiltr. |
| Tm-7 | 100 | 200 | 17 | " | | | | | " |
| Tm-8 | 11 | 90 | 6 | " | | | | | lyophil. |

Most experiments were repeated at least once under identical conditions but for the amounts used for digestion. The code names are defined in III.1.6.

* weight ratio enzyme:peptide;

** amount in mg instead of nmol;

*** incubation with 0.58 M CNBr in 70% formic acid; Tm-7, however, in 70% TFA.

III.1.7.4 Digestion with thermolysin. Thermolytic digestion (6) of demaleylated Tm-7 and [Tm-1b], and of CN-1, CN-3 and CN-8 was carried out in 0.02 M Tris.HCl, pH 8.0, containing 1 mM $CaCl_2$, at a peptide concentration of 0.24-1 µmol/ml, for 2-6 h at 37°C, with 0.03-0.15 mg thermolysin per µmol peptide. The reactions were stopped by lyophilization.

III.1.7.5 Digestion with chymotrypsin of demaleylated Tm-3 and Tm-7 SP-1 was performed for 2 and 3 h, respectively, in 0.1 M $NH_4HCO_3$, pH 8.0, at 37°C and at a peptide concentration of 240 nmol/ml, with 0.12-0.25 mg enzyme per µmol peptide. The digests were lyophilized and the residues were taken up in pH 3.6 electrophoresis buffer.

III.1.7.6 Cleavage with N-Bromosuccinimide (NBS). Demaleylated Tm-7 in 70% acetic acid was reacted, at a peptide concentration of 270 nmol/ml, with 2 mg NBS per µmol peptide, for 1 h at 25°C (103). The reaction was stopped by lyophilization. The residue was taken up in pH 3.6 electrophoresis buffer which resulted in an acid-soluble and an acid-insoluble fraction.

III.1.7.7 Cleavage with CNBr. The demaleylated peptides Tm-5, Tm-7 and Tm-8 were incubated with 1.2, 0.5 and 0.12 ml 0.58 M CNBr, respectively, for 6 or 17 h, at a peptide concentration of 80-200 nmol/ml. Tm-5 and Tm-8 were cleaved in 70% formic acid; Tm-7, in which a Met-Thr bond was assumed to be present, was cleaved in 70% trifluoroacetic acid, a solvent in which peptide bonds between Met-Thr and Met-Ser are supposed to be cleaved to a greater extent than in 70% formic acid. Cleavages were performed under nitrogen in the dark. The resulting solution of Tm-5 was fractionated on a column of Sephadex G-50 sf. The resulting solution of Tm-7 was lyophilized and the residue was taken up in pH 3.6 electrophoresis buffer.

III.1.8 FRACTIONATION AND PURIFICATION OF PEPTIDES

Generally, gelfiltration on Sephadex columns, ion exchange chromatography on anionic and cationic resins (III.1.8.1.) and high-voltage paperelectrophoresis (III.1.8.2.) were used.

III.1.8.1a Gelfiltration on Sephadex columns. In many stages gelfiltration on columns of Sephadex G-50, G-25, G-15 or G-10 was performed, e.g. as the first step in peptide fractionation, for the final purification of peptides or for the removal of excess reagent or salt from peptide solutions.

Usually, columns measured 140-160 x 0.9 cm or 50-60 x 2 cm for mixtures with small amounts of peptides (up to 100 nmol), or 100-140 x 2.3 cm for large-scale separation of amounts up to 3 µmol. Columns were equilibrated and eluted at roomtemperature with the same solvent, which was also used to dissolve the sample ; solvents were 1, 5, 33, or 50% acetic acid, or 0.01 M Tris.HCl, pH ca.8.0, or 1% collidinium acetate, pH 8.3. Flow rates ranged from 5-30 ml/h.

For the first fractionation of the peptide mixtures obtained after cleavage of ([$^{32}$P]-DFP-labeled) atropinesterase with CNBr (fig.V.1 + 2

and fig.VI.5) Sephadex G-50 sf in 5% acetic acid was used. For the first fractionation of peptide mixtures obtained after tryptic digestion of the maleylated ([$^{32}$P]-DFP-labeled) protein (fig.V.5 + 6 and fig.VI.1) Sephadex G-50 sf columns were developed with 0.01 M Tris. HCl, pH 8.0.
Furthermore, Sephadex G-50 sf was used for purification of Tm-8 (fig.VI.3), fractionation of a pool containing Tm-1 (not shown), fractionation of CN-1 and CN-8 (fig. VI.6), fractionation of CNBr- and T-peptides of Tm-5 (fig.V.8 and fig.V.9), purification of one P-peptide of Tm-5 and fractionation of T-peptides of Tm-8 (fig.VII-I.14).
Sephadex G-50 sf was also used for removal of excess maleic acid from [$^{32}$P]-DFP-labeled maleylated atropinesterase and maleylated atropinesterase, for removal of maleic acid after demaleylation, for removal of excess reagent after reduction and carboxymethylation of a pool containing Tm-5 and for desalting of urea-containing peptide fractions.
Sephadex G-25 sf was used for fractionation of CN-3 and CN-10 (fig.V.3), for purification of Tm-5 CNBr-3, for fractionation of T-peptides of Tm-7 (fig. VII-I.11) and for purification of Tm-7 T-3 (fig.VII-I.10).
Sephadex G-15 was used for fractionation of T-peptides of Tm-7 (fig. VII-I.9) and for desalting of the first pool of the tryptic and the clostripain digest obtained after chromatography on Sephadex G-50 sf.
Exact conditions under which fractionation, purification or desalting was performed, such as column size, solvent used, flow rate, fraction size and detection method are given at the experiment in each chapter.

III.1.8.1b Ion-exchange chromatography on anionic or cationic exchange resins. Diethylaminoethyl (DEAE-)cellulose was used for chromatography of the demaleylated and carboxymethylated peptide material containing Tm-5. Experimental details are given in chapter V (V-II.1.2.1).
DEAE-Sephadex A-25 was used for purification of some Tm-peptides (fig.VI.2), while the anion-exchange resin AG 1-X2 was used to fractionate [Tm-]peptides, as described in VII-V.
The cation-exchange resins Aminex A-5, i.e. a purified form of Dowex 50W-X8, and AG 50W-X2 (200-400 mesh) were used for fractionation of digests of Tm- and CN-peptides. They proved to be very effective for the separation of P-peptides of Tm-2, T- and P-peptides of Tm-5, T-, P- and SP-peptides of Tm-6, Ch-peptides of Tm-7 SP-1, P-peptides of Tm-8, p-, sp- and t-peptides of "CN-"peptides, [Tm-]peptides or fragments of [Tm-]peptides and P-peptides of atropinesterase. The columns used, measuring 20-26 x 0.9 cm, were freshly poured with resin regenerated according to Schroeder (115). To regenerate, 20 ml resin (Aminex A-5 or AG 50W-X2) was washed on a Büchner funnel with, successively, 1 l $H_2O$, 0.5 l 0.2 N NaOH, 1.5 l $H_2O$ (pH 7), 0.5 l 0.2 N HCl, 1.5 l $H_2O$ (pH 7), 0.5 l 2M (distilled) pyridine and 0.5 l starting buffer (usually pyridinium acetate). The column was packed at 50°C and washed with at least ten volumes of starting buffer shortly before use, to eliminate ninhydrinpositive material possibly present in the resin. Unless otherwise stated, samples were applied to the column in a

pyridinium acetate buffer with the same pyridine concentration as the starting buffer, but with a 0.2-0.5 pH-units lower pH. Tryptic and peptic digests of Tm-5 were exceptional, as these were applied as a solution in 60% acetic acid and a few drops formic acid or dissolved in 40% acetic acid. Before samples were applied to an Aminex column an aliquot was fractionated by paperelectrophoresis at pH 3.6 to estimate the number of fragments and to get an impression of their polarity on the basis of their electrophoretic behaviour. Peptides migrating on paper to the negative electrode were believed to possess an appropriate size and polarity to be eluted from Aminex.

Cation-exchange columns were developed with the buffers given in table III.2; during elution a gradient in pH and pyridine concentration was applied. Pyridine and other volatile bases were distilled over ninhydrin before use. The buffer molarities quoted refer to pyridine. Elution, with a flow rate of about 12 ml/h, was initiated with a small volume of the starting buffer, followed by one or two gradients. Sometimes development was terminated with 2M pyridine to remove strongly adsorbed material. When higher flow rates were applied poorer separation was obtained as was found for P-peptides of atropinesterase. Fractionation of peptides obtained after digestion of Tm-peptides was usually performed with a two-steps gradient, the first going from 150 à 200 ml 0.05 M pyridinium acetate pH 2.4 (C) to 0.2 M pyridinium acetate pH 3.1 (D) and the second going from 250 à 500 ml 0.2 M pyridinium acetate pH 3.1 (D) to 2M pyridinium acetate pH 5.0 (E).

Table III.2 Composition of the aqueous buffers used for development of columns of Aminex A-5, AG 50W-X2 and AG 1-X2.

| Buffer | | pH | Ingredients (ml/l) | | |
|---|---|---|---|---|---|
| | | | acetic acid | pyridine | second base |
| pyridinium acetate | molarity* | | | | |
| A | 25 mM | 1.6 | 600 | 2 | |
| B | 25 " | 1.9 | 600 | 4 | |
| C | 50 " | 2.4 | 340 | 4 | |
| C.2 | 0.1 M | 3.0 | 200 | 8 | |
| D ** | 0.2 " | 3.1 | 278 | 16.3 | |
| D.2 0.5 | " | 3.8 | 190 | 40 | |
| E ** | 2.0 " | 5.0 | 142.5 | 162.5 | |
| F | 2.0 " | | - | 162.5 | |
| 2-picolinium/ pyridinium acetate | | | | | |
| G ** | | 8.0 | ca.0.1 | 11.8 | 28.3 |
| 2,4,6-collidinium/ pyridinium acetate | | | | | |
| H ** | | 8.3 | 0.4 | 10 | 10 |
| N-ethylmorpholinium/ pyridinium acetate | | | | | |
| K ** | | 9.3 | 0.1-0.5 | 7.5 | 12.5 |

* The molarities refer to pyridine.
** Buffers described by Schroeder (115).

Fractionation of fragments of "CN-"peptides was performed with a gradient going from 0.1 M pyridinium acetate pH 3.0 (C.2) to 0.5 M pyridinium acetate pH 3.8 (D.2) followed by a second from 0.5 M pyridinium acetate pH 3.8 (D.2) to 2 M pyridinium acetate pH 5.0 (E).
Preparation of gradients. Gradients were made with a multichannel peristaltic pump by pumping the second buffer into a mixing chamber containing the first buffer at a flow rate which was approximately half (45%-50%) the flow rate used to elute the column. In an early stage of this study gradients were made by use of a siphon. The eluent was pumped out of the mixing chamber filled with the first buffer, which was connected by a siphon to a second vessel containing an equal volume of the second buffer. The first mentioned method was technically more reliable and was used in most chromatographic procedures.

III.1.8.2 High-voltage paperelectrophoresis (HVE). Peptide separation was performed by HVE at pH 3.6 or 6.5 on chromatographic paper immersed in tanks with cooled varsol (84) at 50 V/cm for 0.5-3 h (high voltage equipment, type HV-5000B-500, Savant Instruments, Inc., Hicksville, N.Y., U.S.A). Usually the paper was moistened with pyridine-acetic acid-water (1:10:89, by volume), pH 3.6, or when peptides were poorly resolved at pH 3.6, with pyridine-acetic acid-water (25:1:225, by volume), pH 6.5, as described by Ambler (3). Peptides were lyophilized, dissolved in the appropriate buffer and applied on Whatman no.1 paper (length 80 cm, width varying with the amount to be separated between 10 and 25 cm) as a band of 1-8 cm with loads of 10-50 nmol/cm. During this application, drying was facilitated by the use of an unheated hair dryer. For the pH 3.6 buffer system the sample was applied 12 cm from the bottom (anodic) edge of the paper because at this pH all peptides were positively charged or neutral. When the buffer with pH 6.5 was used the sample was applied in the middle of the paper. After electrophoresis the paper was dried in a ventilated hood at roomtemperature. Peptides were located by staining guidestrips or the whole paper (III.1.9.2). For their recovery, zones with peptide material were cut out and eluted with pH 3.6 buffer by a form of descending chromatography. To this end, paperstrips were fixed between 2 glass rods in a closed chamber, with the upper tip placed in a trough containing the eluent. The effluents were collected in vials and stored at -20 $^{o}$C.

III.1.9 DETECTION OF PROTEINS AND PEPTIDES

Detection of protein and peptides in column effluents was carried out directly, by measuring the ultraviolet (UV) absorbance at one or more wave-lengths or by determination of the radioactivity, if present, or after colour reaction with ninhydrin. Which method(s) was (were) used depended on the properties of the peptide, the amount of peptide material and on the solvent used for column development. To detect peptides on paper several staining methods, general or specific, were employed.

III.1.9.1 Detection of protein or peptides in column effluents.

III.1.9.1a Measurement of the UV-absorbance. Depending on the composition of the peptide(s), the absorbance at 280, 250 and/or 225 nm was determined in a Zeiss M4 Q III spectrophotometer.

III.1.9.1b Determination of the $^{14}C$ and/or $^{32}P$ radioactivity. In general this was done by liquid scintillation counting (LSC) of small aliquots (0.5-3%) of all relevant fractions; 0.02-0.4 ml samples were mixed in glass vials with 13 ml scintillation fluid (initially toluene-triton X-100 (3:1, by volume), later tritosol (52)) and counted in a Nuclear Chicago Mark II, 3 or 4 liquid scintillation spectrometer. Occasionally, $^{32}P$-labeled peptides were detected during development of the column by elution via a spiralized narrow plastic capillary placed in front of an end-window counting tube; the measured radioactivity was registered on the multichannel recorder of the peptide analyzer as shown in fig.V.1A and V.5.

III.1.9.1c Detection with ninhydrin. By use of a home-made peptide analyzer (PA), made according to the principle described by Catravas (25), peptides were detected automatically by measurement of the absorbance at 570 nm (A570) after partial alkaline hydrolysis and reaction with ninhydrin. Alkaline hydrolysis, neutralization, colour reaction with ninhydrin and measurement of the A570 were performed automatically according to Hirs (64) as modified by Catravas (25) and Hill and Delaney (63). The heart of the PA was the very reliable 12 channel peristaltic pump (Cenco Instrumenten, Breda, The Netherlands); the absorbance was measured in a spectrophotometer, equiped with a flow-cell (light path 2 cm) and a 570 nm filter, also from Cenco. This method was used preferably throughout this study, because all peptides present in the effluent could be detected, irrespective of their amino acid composition, and because the method could be used to locate very small amounts of peptide (up to 5 nmol) without consuming most of it, owing to the fact that peak-broadening and sample-uptake were kept to a minimum. Analysis was carried out either:
a) on line during chromatography, by feeding a proportion of the effluent directly into the PA at a flow rate of about 1.8 or 3.0 ml/h.
b) discontinuously after chromatography, by feeding aliquots (about 1%) of each fraction into the PA, for 10-60 seconds, at a flow rate of 1.8 or 3.0 ml/h, alternated with the uptake of the solvent used for elution. Sample change was performed manually or by use of an automatic sample changer (Cenco Instrumenten). The method of choice (a or b) depended on the amount of each peptide present in the peptide mixture and the flow rate used to develop the column. In all figures the A570 is given in recorder-units which are only a rough estimate of the A570; furthermore the A570 is not strictly proportional to the total amount of amino acids in a peak since hydrolysis of peptide bonds occurred to different extent depending on the length of the polypeptide chain.

III.1.9.2 Detection of peptides on paper. Peptides were detected on paper with the general staining reagents fluorescamine and ninhydrin, while specific staining methods were applied to find His-, Tyr-, Trp- and Arg-containing peptides.

III.1.9.2.1a Staining with fluorescamine. Guide strips or the whole paper were dipped into a solution of fluorescamine (130) prepared by mixing 1 mg fluorescamine in 20 ml aceton (A) with an equal volume of 1% pyridine in aceton (B). When the whole paper was used for detection, solution A contained 0.1 mg fluorescamine. After drying in the air, peptides became visible as fluorescent spots under a "long wavelength UV-lamp", at 366 nm.

III.1.9.2.1b Staining with ninhydrin. Staining with ninhydrin was performed in different ways: either the paper was sprayed with a solution of 0.2 g ninhydrin in 87 ml n-butanol, 4 ml acetic acid and 9 ml water (reagent I) and heated in an oven at 70°C for 10-15 min to develop the bluish-purple colour; or it was stained with a cadmium-ninhydrin reagent (II) by dipping the paper into a solution containing 0.85 g ninhydrin in 85 ml aceton, 0.1 g cadmium acetate in 5 ml glacial acetic acid and 10 ml water, followed by a period of about 18 h at roomtemperature or of a few minutes in a stove at 70°C needed for the red or orange colour to develop (41).
A modified cadmium-ninhydrin stain (III) was used as described by Yamada and Itano (137) when the paper had been stained with the Arg-reagent of these authors. Then, the reagent consisted of 1 g ninhydrin, 80 ml aceton, 15 ml glacial acetic acid and 0.1 g cadmium acetate dissolved in 5 ml water. The high concentration of acetic acid was necessary to neutralize the alkalinity of the Arg-reagent (see below).

III.1.9.2.2a Arginine-containing peptides were detected according to Yamada and Itano (137). The paper was dipped into a solution prepared by mixing equal volumes of 0.02% phenanthrenequinone (w/v) in absolute ethanol and 10% NaOH (w/v) in 60% (v/v) ethanol (both freshly prepared). After drying in the air for 15 min, Arg-containing peptides were visible under a long wave-length UV-lamp, at 366 nm, as greenish or bluish fluorescent spots against a dark blue background.

III.1.9.2.2b Histidine- and tyrosine-containing peptides were detected with the Pauly reagent as described by Easley (46) with slight modification: to a freshly prepared solution of 0.5 g sulphanilic acid in 50 ml 1 N HCl, chilled in an ice bath, 2.5 g $NaNO_2$ was added; after 0.5 h at 0°C the paper was dipped into the icecold solution. His-containing peptides became visible as cherry-red spots, Tyr-containing peptides as reddish-brown spots.

III.1.9.2.2c Tryptophan-containing peptides were detected with the Ehrlich reagent by dipping the paper into a solution consisting of 90 ml aceton, 10 ml concentrated HCl and 1.0 g dimethylaminobenzaldehyde (10). Trp-containing peptides developed a purple colour after a few minutes at room temperature.

III.1.9.2.3 Combination of staining techniques. Usually, both fluorescamine- and ninhydrin staining were applied, in that order, because it had been shown that peptides with an N-terminal Asp, Asn, Glu, Pro or pyroglutamyl residue (88) gave low fluorescence with fluorescamine despite the presence of internal or C-terminal Lys-residues; free acidic amino acids, too, were poorly detected (124). These peptides (peptides with N-terminal Pro and pyroglutamine only when containing an internal or C-terminal Lys-residue) and free amino acids gave clear spots with the ninhydrin reagent.
Specific staining techniques were applied after fluorescamine treatment, in the sequence recommended by Easley (46): staining for Arg came before ninhydrin (III) or His-staining; staining for Trp followed after ninhydrin (I). His-staining was performed without subsequent ninhydrin detection.

III.1.10 SEQUENCE DETERMINATION
Amino acids were released one at a time from the N-terminal side of the peptide or protein by the Edman degradation procedure (47). After each step the new N-terminal residue was identified, usually as the dansyl-derivative (56), but sometimes as the phenylthiohydantoin amino acid. Beside the manual Edman degradation procedure use was made of the automated sequencing method in a "spinning cup" sequencer.

III.1.10.1 Manual dansyl-Edman degradation and identification of dansyl (DNS)-amino acids. Most of the amino acid sequences of the peptides were established by the manual dansyl-Edman method as described by Gray for extended degradation (59) with a modification of the coupling procedure as described by Peterson (100).

III.1.10.1a Degradation procedure. The peptide (5-50 nmol) was dissolved in a small volume of "coupling buffer", i.e. 0.4 M N,N-dimethyl-N-allylamine in 1-propanol-water (3:2 by volume) titrated to pH 9.5 with trifluoroacetic acid (TFA) and an aliquot was taken for dansylation (the volume of coupling buffer was chosen such that 20 µl solution or less would suffice for dansylation, in order to reduce introduction of undesired impurities). Subsequently, more coupling buffer was added to a volume of 200 µl, 20 µl phenyl isothiocyanate (PITC) was added and left to react for 30 min at 50°C. (With peptides containing more than 5 residues, after each third or fourth cycle the reaction mixture was extracted in this stage with 0.5 ml benzene). After removal of the solvents in vacuo at 50°C for 15 min, cleavage was performed by treatment with 100 µl anhydrous TFA at 50°C for 10 min. A purple solution during incubation with TFA generally indicated the presence of N-terminal Trp, a residue which can be missed easily during identification. When Asn was expected to occur at the N-terminus, the incubation time with TFA was brought back to 5 min, to prevent β-rearrangement. After removal of TFA in vacuo at 50°C, the residue was dissolved in 150 µl water and extracted twice with 0.6 ml ethyl acetate. After the resulting aqueous phase had been taken to dryness in vacuo at 50°C, the shortened peptide was ready for the next

degradation cycle. All additions and incubations were carried out under nitrogen or argon to prevent side-reactions caused by oxygen. Occasionally hydrophobic peptides (e.g. Tm-7 and Tm-3 ) with internal Lys-residues were coupled with sodium 4-sulfophenyl isothiocyanate (Braunitzer's reagent) prior to the first cycle (18), to introduce a negative charge at the $\varepsilon$-$NH_2$-groups, which otherwise would become blocked by reaction with the hydrophobic PITC; in the latter case, the enhanced hydrophobicity might lead to removal of the peptide by the extraction with ethyl acetate. Coupling was performed in 100 µl 0.4 M N,N-dimethyl-N-allylamine buffer, pH 9.5, containing 0.5 mg 4-sulfophenyl isothiocyanate, for 2 h at 54°C. After lyophilization the 4-sulfophenylthiocarbamyl peptide was subjected to the first step of the normal PITC coupling procedure.
Sometimes "ragged" degradation was observed, i.e. the contamination of the newly formed DNS-amino acid with derivatives of previous steps. This phenomenon frequently started when a Gly- or Pro-residue became the N-terminal residue of the foreshortened peptide.

III.1.10.1b Identification of N-terminal residues with dansylchloride. Dansylation of peptides was carried out by dissolving 0.5-10 nmol lyophilized peptide in 20 µl 0.05 M $NaHCO_3$ (pH adjusted to 10.0 with 0.1 N NaOH); after drying in vacuo, the peptide was dissolved again in 20 µl 0.05 M $NaHCO_3$, pH 10.0 (to be sure that the pH would be high enough for dansylation), mixed with an equal volume of dansylchloride in aceton (5 mg/ml), incubated for 15 min at 50°C, dried in vacuo and hydrolyzed with 50 µl 6 N HCl in a sealed Pyrex tube, for 6 h at 110°C unless otherwise indicated. When DNS-Trp was expected, hydrolysis was performed with 50 µl 6 N HCl containing 2% TGA. When DNS-Val or DNS-Ile were expected the hydrolysis time was extended to 18 h or longer; in case of DNS-Pro it was shortened to 4 h. The hydrolysate was dried in vacuo (drying was particularly thorough when TGA had been added, which compound disturbs identification of DNS-amino acids). The residue was dissolved in 5 µl ethanol; 2 µl of the clear solution was applicated on a polyamide sheet (5 x 5 cm) according to Woods and Wang (135). Sheets were developed according to the method of Hartley (60) by ascending chromatography in two directions, with three or four successive solvents; first direction, solvent 1 : 1.5% (v/v) formic acid), second direction, solvents 2, 3, 4; solvent 2: initially benzene-acetic acid (9:1, by volume) later on substituted by toluene-acetic acid (10:1, by volume); solvent 3: ethyl acetate-methanol-acetic acid (20:1:1, by volume); solvent 4: pyridine-acetic acid-water-ethanol (9:16:1000:342, by volume). DNS-amino acids were detected under 254 nm UV as fluorescent spots, which were photographed. All amino acids were separated with 3 solvents, except the basic amino acids which separated during development with solvent 4.
The positions on the chromatogram of the DNS-amino acids were as described by Hartley (62). DNS-Trp was located as described by Giglio (55) near the position of bis-DNS-Lys. DNS-CMCys was located close to DNS-Asp. The position of DNS-Hse, DNS-Hse-lactone and $\alpha$-DNS-$\varepsilon$-phenylthiocarbamyl-Lys ($\alpha$-DNS-$\varepsilon$-PTC-Lys) was as described by Scheffer (112).

Identification of internal Lys-residues caused some problems, as α-DNS-ε-PTC-Lys was found at the position of Leu after chromatography with 2 solvents and moved to the position of the Phe in the third solvent and, generally, spots were very faint. In contrast to these difficulties with Lys, the identification of α-DNS-Lys in a maleylated peptide was performed very easily.
In some cases DNS-dipeptides were detected: DNS-Pro-Val was found between Pro and the second solvent-front (solvent 1+2), while DNS-Ile-Ile was found at the position of Phe after chromatography with 2 solvents and moved to the position of Leu after chromatography with 3 solvents (112). Also DNS-Val-Leu coincided with Phe after 2 solvents and with Leu after chromatography with the third solvent, while DNS-Val-Tyr was found at the position of Tyr after chromatography with 2 and 3 solvents. Usually, the correct DNS-amino acids could be obtained by prolongation of the hydrolysis time up to 24-48h.

III.1.10.2 Isolation and identification of phenylthiohydantoin derivatives (PTH-derivatives) after manual Edman degradation. When confirmation of amino acids identified as DNS-residues was intended or when Asn had to be distinguished from Asp or Gln from Glu, a modified procedure for the Edman degradation had to be applied, as described in III.1.10.2a.

III.1.10.2a Modified degradation procedure and isolation of PTH-amino acids. At the end of the coupling step (see III.1.10.1a) the reaction mixture was extracted twice with 0.5 ml benzene (added, centrifuged and removed under nitrogen) and the extract was discarded. After the cleavage with TFA and extraction with ethyl acetate, the extract containing the thiazolinone derivative to be identified was obtained. The solvent was evaporated under a stream of nitrogen and the derivative was converted into the correspondig PTH-amino acid by heating for 10 min at 80°C with 0.2 ml 1 N HCl, according to Edman and Begg (49). Then the mixture was extracted three times with 0.8 ml ethyl acetate. The pooled extracts were evaporated under nitrogen and the residue containing the PTH-amino acid was dissolved in 5 µl ethyl acetate.

III.1.10.2b Identification of PTH-amino acids was performed either by thin layer chromatography or by amino acid analysis after "back hydrolysis" to the parent amino acid.
To identify the PTH-derivative by thin layer chromatography, the ethyl acetate solution was spotted onto a silica gel coated plastic thin layer sheet (6.3 x 6.3 cm) with fluorescence indicator, according to Cohen Solal et al. (32). After ascending chromatography in chloroform-methanol (9:1, by volume) (20) sheets were dried for 5 min at 110°C and PTH-amino acids were detected and identified according to their position by examination under UV at 254 nm.
Additional identification: when more than 5 nmol PTH amino acid was expected to be present, the method described by Roseau et al. (109) was carried out: prior to sample application, sheets were activated by

heating for 15 min at 110°C; after chromatography and examination under UV, sheets were sprayed with 0.1% (w/v) ninhydrin in a solution of 95% (v/v) ethanol and 5% (v/v) collidine and heated at 110°C. Within 5 to 15 min PTH-amino acids present in sufficient amount became visible as coloured spots. The PTH-derivatives of Asn, Asp, Gln and Glu showed specific colours which facilitated their identification. When less than 5 nmol PTH amino acid was present, the staining method of Boigne et al. (16) was used as applied by Cohen Solal et al. (32): prior to sample application the layer was soaked in a 1% (w/v) solution of soluble starch and dried for 30 min at 80°C. After chromatography, drying at 110°C and examination under UV, the sheet was dipped into a solution containing 15 ml 10% sodiumazide, 10 ml 0.1 N $I_2$ solubilized with KI (about 1 g/10 ml) and 25 ml water. PTH-amino acids appeared immediately as white spots on a brownish background. As little as 0.5 nmol of PTH-amino acid could be detected by this staining procedure.

Sometimes, PTH-amino acids were identified and quantitatively determined after back hydrolysis of the PTH-amino acid to the parent amino acid. For this purpose the dry residue was incubated with 0.1 ml 57% HI at 130°C for 18 h (120); after lyophilization, the free amino acids (or derivatives) were identified by amino acid analysis. CMCys was determined after back hydrolysis of the PTH-derivative with 57% HI, giving a product which during amino acid analysis is eluted at the position of Ala, as shown by reaction of standard CMCys with PITC, subsequent conversion with HI and amino acid analysis.

III.1.10.3a Automatic Edman degradation was carried out according to Edman and Begg (49) in a Beckman 890 C (ser.nr.511) sequencer equipped with a Balzer OFM 002 oilfilter. The Beckmanprogram nr. 122974 was used, adapted with respect to Quadrol-flow because 0.25 M in stead of 1 M Quadrol was used, while the fraction collector was not kept under vacuum. Polybrene (ref. 128, 72; 3 mg per run), a polymeric quaternary ammonium salt, was used as a carrier to retain the peptide in the spinning cup. Chemicals were purified as described by Frank (50). Conversion of the anilinothiazolinones to PTH-amino acids was performed according to the method of Amons et al.(private communication): the chlorobutane extracts containing the thiazolinone derivatives to be identified, were collected in the fraction collector, in tubes containing 0.3 ml 1 N HCl and 3 µl ethanethiol. The chlorobutanes were blown off with a stream of nitrogen at 30°C. Tubes were stoppered and the derivatives were converted into the correspondig PTH-amino acids by heating for 10 min at 80°C. After lyophilization or evaporation on a Rotavapor the residues were dissolved in 0.1 N sodium acetate pH 5.2, containing 36% acetonitrile.

III.1.10.3b Identification of N-terminal residues as PTH-derivatives. An aliquot of the acetonitrile solution (see preceding section) was analyzed by HPLC, essentially by the method of Zimmerman (139) and Schlesinger (113) as modified by Amons (personal communication). Analysis was performed on a Spherisorb 5 ODS column (25 x 0.46 cm) preceded by a Co Pell ODS column (10 x 0.21 cm). The

larger column was housed in an oven thermostatted at 62°C, the smaller column was operated at roomtemperature. Columns were equilibrated and eluted with 0.01 N sodium acetate pH 5.20, containing 36% acetonitrile. The equipment used was a Pye-Unicam LC 3 chromatographic system consisting of a pump XP, a UV-detector measuring the absorbance at 269 nm, a Valco 7000 injector and a Kipp en Zonen type BD 8 strip chart recorder. All PTH-amino acids were separated in 13 min, by using a flow rate of 1 ml/min and a backpressure of about 62 ato.
Sometimes, the remaining part of the solution containing the PTH-amino acid was treated with 57% HI for the conversion into the free amino acid (see III.10.2b for further details).

III.1.11 AMIDE ASSIGNMENT
The question which Glu- and Asp-residues found after acidic hydrolysis originated from glutaminyl and asparaginyl residues in peptides was answered (i) by digestion with LAP or CPA and CPB (III.1.5.1a; III.1.5.2) followed by amino acid analysis or (ii) by identification of the PTH-amino acids, obtained after direct Edman degradation, with thin layer chromatography (III.1. 10.2b), or (iii) by determination of the electrophoretic mobilities of peptides at pH 6.5, according to Offord (93). Additional information was provided by the specificity of S.protease.

III.1.12 METHANOLYSIS OF PYROGLUTAMIC ACID
Opening of the pyrrolidone ring of pyroglutamic acid was achieved as described by Kawasaki and Itano (70). Details are given in chapter VI-II.2.1.

III.2 MATERIALS
Atropinesterase from Pseudomonas putida bacteria (biotype A, strain L PMBL-1) was isolated and purified in our laboratory by Oosterbaan et al. (to be published). The specific activity of the preparations used varied between 500 and 600 U per A280.
Proteolytic enzymes. Trypsin from porcine pancreas (3x crystallized; batch 7016) was obtained from Miles Laboratories (Slough, U.K.); trypsin from bovine pancreas (treated with; 2.0 U/mg) and thermolysin (crystallized; 6000 PU/mg) were from Merck A.G. (Darmstadt, Germany); α-chymotrypsin ( 3x crystallized; C-4129, lot no. 40 F-805 I; 48 U/mg) was from Sigma Chemical Company (St.Louis, U.S.A.); Staphylococcus aureus V8 proteinase (batch 276; 480 U/mg) was from Miles Laboratories (Slough, U.K.); pepsin (3x crystallized lot no. 5601) was from Nutritional Biochemicals Corporation (NBC) (Cleveland, Ohio, U.S.A.); carboxypeptidase A and carboxypeptidase B (treated with DFP) were from Boehringer (Mannheim, Germany), from Sigma Chemical Company (lot no. 127C-8055; 55 U/mg and lot no. 119C-80751; 180 U/mg) or from Worthington Biochemical Corp. (Freehold, N.J., U.S.A.). Leucinaminopeptidase was from Worthington (lot no. 55B 411; 100 U/mg) or from Sigma (lot no. 50F-8096; 18.2 U/mg). Of some enzyme preparations several brands were used because some samples delivered by one company or the other showed inconvenient aspecific actions.

Materials for chromatography and electrophoresis.

Sephadex G-50, G-25, G-15 and G-10, fine and superfine grade and DEAE-Sephadex A-25 were obtained from Pharmacia (Uppsala, Sweden). Aminex A-5 and AG 50 W-X2 (200-400 mesh) and AG 1-X2 (200-325 mesh) were obtained from Bio-Rad Laboratories (Richmond, California, U.S.A.). DEAE-cellulose DE-52 and Whatman no.1 paper were obtained from W. and R.Balston (Maidstone, Kent, U.S.A.). Polyamide thin-layer sheets (F 1700) were from Schleicher and Schüll (Dassel, Germany). Plastic thin-layer chromatography sheets, silica gel 60 F 254, with fluorescence indicator were from Merck.

Reagents.

(-)-Atropine (hyoscyamine) sulfate was obtained from Nutritional Biochemicals Corporation. [$^{32}$P]-DFP was synthesized by the chemical laboratory of RVO-TNO Rijswijk (Z-H), The Netherlands. Maleic anhydride (chemical pure) and cadmium acetate were from the British Drug House Ltd. (Poole, U.K.). Maleic anhydride was purified by sublimation before use. Iodoacetic acid and 2-mercaptoethanol were from Merck. Iodoacetic acid was recrystallized from petroleumether 40-60. Urea was from UCB N.V. (Brussels, Belgium). EDTA was from Siegfried S.A. (Zofingen, Switzerland). [$^{14}$C]-maleic anhydride and 2-[$^{14}$C]-iodoacetic acid were obtained from The Radiochemical Centre (Amersham, U.K.). N-Bromosuccinimide was from Koch-Light Laboratories Ltd.(Colnbrook, Buchs, U.K.). Cyanogen bromide and dansylchloride were from Fluka A.G. (Buchs, Switzerland). Fluorescamine was from Hoffman-la Roche (Basel, Switzerland). Ninhydrin was a product of Pierce Chemical Corporation (Rockford, Illinois, U.S.A.). Phenanthrenequinone-(9,10) was purchased from Merck. Pyridine, 2-picoline and sym.-collidine (2,4,6-trimethylpyridine) and N-ethylmorpholine, were purchased from Merck; these solvents used for column chromatography, were distilled over ninhydrine twice before use (10 g ninhydrin/l). Acetic acid (analytical grade) was from Merck. Trifluoroacetic acid (Merck) was distilled 3 times after refluxing with chroomtrioxide ($CrO_3$), dried over CaSO and stored at 4$^{o}$C in sealed ampoules. N,N-dimethyl-N-allylamine and other reagents used for Edman degradation such as pyridine, benzene, ethyl acetate, phenyl isothiocyanate, sodium-monohydrate of 4-sulfophenyl-isothiocyanate and n-propanol, and the standard PTH-amino acids were sequenal grade products from Pierce. N,N-dimethyl-N-allylamine was distilled over KOH and pthalic anhydride and stored at 4$^{o}$C in sealed ampoules. Reference samples for the amino acid analyzer, containing equal molar amounts of all amino acids were obtained from Organon (Oss, The Netherlands).

# CHAPTER IV CHARACTERIZATION OF ATROPINESTERASE

## INTRODUCTION

The atropinesterase used in these studies is produced by Pseudomonas putida bacteria (biotype A, strain PMBL-1) as an intracellular enzyme. It was isolated by Rörsch and Berends and characterized as an esterase because of its ability to catalyze specifically the hydrolysis of (−)-atropine into the defined products tropine and tropic acid (108). Because of this property and because the activity is inhibited by the organophosphorus compounds Soman, DFP and Sarin, the enzyme was assumed to belong to the "serine-esterases", a group of hydrolytic enzymes that contain in their active center one serine residue with an activated hydroxyl group. Furthermore it had been shown that the enzyme was stable below 40°C and within a pH range of 5-10 (126).
The stability and further properties allowed isolation in good yields, so that studies could be performed to elucidate the primary structure of this new representant of serine-enzymes. The large quantities of enzyme needed for this enterprise were isolated in our institute by Oosterbaan and Kouwenberg (Oosterbaan et al., to be published). They took care of the large-scale growth of Pseudomonas putida bacteria and the isolation and purification of atropinesterase.
In this chapter some properties of atropinesterase are discussed. The degree of inhibition with [$^{32}$P]-diisopropyl phosphorofluoridate (DFP) is described in relation to $^{32}$P-binding, studied to establish the specificity of the reaction and to determine the number of active sites per amount of active enzyme. Experiments performed to assess the purity of the final enzyme preparations used for sequence studies are presented, as well as the determination of some characteristics of the enzyme as a protein, i.e. the molecular weight, the amino acid composition and part of the N-terminal and C-terminal sequence.

## IV.1 INHIBITION OF ATROPINESTERASE WITH [$^{32}$P]-DFP AND DETERMINATION OF THE CATALYTIC CENTER ACTIVITY

The specificity of the reaction with [$^{32}$P]-DFP was studied; during incubation of atropinesterase (preparation 1) with [$^{32}$P]-DFP, samples were taken at intervals in which the percentage of inhibition and the amount of $^{32}$P bound to the protein were determined. The results, plotted in fig.IV.1, showed a linear relationship between the two determined phenomena. Evidently the amount of $^{32}$P bound per inhibited unit of enzymatic activity is practically independent of the degree of inhibition or the duration of the incubation, as can be seen from the figures given in table IV.1 (mean value 33.6 pmol $^{32}$P per inhibited unit). A comparable effect was demonstrated by Rörsch et al. (108) during incubation with Soman (1,2,2-trimethylpropyl methylphosphonofluoridate) and confirmed by Oosterbaan et al. (97). This linear relationship together with the observation that the $^{32}$P incorporated did not exceed a certain amount after prolonged incubation (up to 18 h), indicates that the reaction with DFP -just as with Soman- exclusively occurs at the active site, at least under the conditions of the experiments.

From the data the catalytic center activity was calculated, i.e. the ratio between the enzymatic activity of a solution under the conditions used for the standard activity determinations and the concentra-

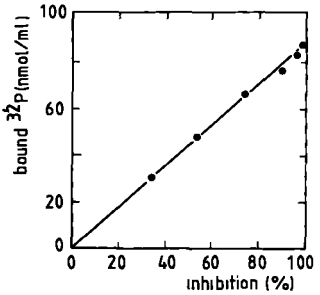



Fig.IV.1. Inhibition of atropinesterase with [$^{32}$P]-DFP.
The active site concentration of the sample used (preparation 1) was ca.90 nmol/ml; inhibition with 0.61mM [$^{32}$P]-DFP was followed during 4 h at roomtemperature; at time intervals samples were taken in which the percentage of inhibition and the amount of $^{32}$P bound to the protein were determined (see also IV.2.1).

Table IV.1 Reaction of atropinesterase with [$^{32}$P]-DFP: Time course of inhibition and of $^{32}$P-binding.

| Time (min) | inhibited enzyme activity (%) | $^{32}$P bound to protein (nmol P/ml) | Ratio $^{32}$P-binding (nmol/ml) / inhibition (%) | Catalytic center activity (*) |
|---|---|---|---|---|
| 0 | 0 | 0 | - | - |
| 18 | 33.4 | 31.3 | 0.937 | 28,700 |
| 32 | 52.8 | 48.4 | 0.917 | 29,400 |
| 62 | 73.3 | 67.0 | 0.914 | 29,400 |
| 116 | 89.2 | 77.2 | 0.865 | 31,100 |
| 182 | 95.1 | 84.0 | 0.883 | 30,500 |
| 248 | 97.3 | 88.0 | 0.904 | 29,800 |
| | | | mean 0.903 | 29,800 |

During incubation of atropinesterase, 2691 U/ml, with [$^{32}$P]-DFP at intervals enzyme activity and $^{32}$P-binding to protein were measured.

* Number of substrate molecules hydrolysed per min per active center, i.e. enzyme activity units per µmol catalytic center. Calculation: per ml 1% inhibition corresponds to 26.9 U and about 0.903 nmol phosphorylated active sites.

tion of the active centers. The catalytic center activity was also determined after complete inhibition of several other preparations (nrs. 2 through 6). These values, together with the value of preparation 1 gave an average activity of 29,800 $\pm$ 1,500 molecules (−)-atropine hydrolysed per minute per catalytic center, which agrees very well with the value of 30,700 found after inhibition of atropinesterase with [$^{32}$P]-Soman by Oosterbaan et al. (97). These results indicate that atropinesterase behaves as all other serine-hydrolases studied in this respect: covalent binding of the usual organophosphate inhibitors occurs exclusively at the active serine, irrespective of the exact chemical structure of the inhibitor or the duration of the incubation.

IV.1.1 Experimental details

Preparation 1: 5.2 ml atropinesterase (2691 U/ml; SA 534) in 0.03 M Tris.HCl, pH 8.2, were incubated with [$^{32}$P]-DFP at a DFP concentration of 0.61 mM. Preparations 2 through 6 were (almost) completely inhibited with [$^{32}$P]-DFP under identical conditions (III.1.1.3). At intervals during the incubation (prep. 1) and after complete inhibition (preps. 1 through 6) samples were taken in which the remaining enzymatic activity (III.1.1.1) and the amount of $^{32}$P bound to the protein (III.1.1.3) were determined. To determine the catalytic center activity the inhibited enzymatic activity was related to the amount of protein-bound $^{32}$P.

## IV.2 ISOLATION AND PURIFICATION OF ATROPINESTERASE

To obtain large quantities of (almost) pure atropinesterase, kilogram quantities of pseudomonas bacteria were cultured. Essentially the method described by Rörsch et al. was used, but several modifications were introduced in the extraction of atropinesterase from the cells in order to increase the yield, and also in the purification of the enzyme, so that pure atropinesterase was obtained in good yield (ca.15%) after the final purification step. (In the procedure followed, from 3 kg bacteria 100 mg pure enzyme was recovered.) The purity of these preparations was studied by polyacrylamide gelelectrophoresis, isoelectric focusing and endgroup determinations.

### IV.2.1 CHECKING OF THE PURITY OF ATROPINESTERASE PREPARATIONS

The purity of the preparations of atropinesterase used for sequence studies (specific activity (SA) between 500 and 600) was checked by polyacrylamide gelelectrophoresis with and without SDS (III.1.2.1), by isoelectric focusing (III.1.2.2) and by endgroup determinations (III.1.5).

Polyacrylamide gelelectrophoresis (PAGE) without SDS was performed on samples (ca. 90 U) in Tris.HCl buffer in 7% polyacrylamide gels at pH 8.3 and gel electrophoresis with SDS was performed on samples which were pretreated with 1% SDS, 4 M urea and 1% 2-mercaptoethanol, in 10% gels at pH 7.2 (44, 134). After electrophoresis and staining with Coomassie Brilliant Blue R250, one deeply coloured protein band was observed, while in a few preparations one additional faint band of contaminating protein material was detected, as is shown in fig.IV.2. These results indicated that protein preparations with a SA between 500 and 600 contained a pure or almost pure protein.

Isoelectric focusing (IEF) of samples of purified atropinesterase (ca.30 U) was performed on polyacrylamide sheets with Ampholites from pH 3.5 to 9.5. After staining with Coomassie Brilliant Blue R250 one very strong band and some faint bands were generally observed, an example of which is given in fig.IV.2. In this photograph beside the strong band at pH 5.7 several other faint bands can be detected at lower pH values. Most of these faint protein zones contained active

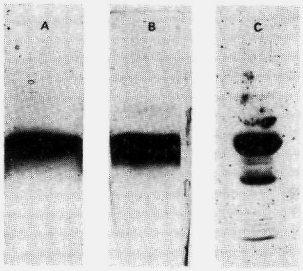



Fig.IV.2. Polyacrylamide gelelectrophoresis and isoelectric focusing of atropinesterase. Gelelectrophoresis (fig.IV.3 A and B) of atropinesterase on 7% gels at pH 8.3 and isoelectric focusing with ampholytes in the range of pH 3.5 to 9.5 (fig.IV.3 C). The amounts applied to the gels in fig.IV.3 A, B and C were 50, 90 and 25 units of enzymatic activity, respectively. In fig.IV.3 B and C the same preparations were used. Protein zones of fig.IV.3 C were shown to be enzymatically active.

atropinesterase as was revealed by a specific colour reaction. These results indicated that the ratio between active enzyme and contaminating proteins was much better than was assumed after protein staining and that the enzyme preparations used for sequence studies were (almost) pure. The zones of atropinesterase with a lower isoelectric point than the main band could be caused by conversion of amides into acidic residues during one or more steps of the purification procedure.

Endgroup determinations at the N-terminal side of the protein by dansylation and by application of the automated Edman procedure confirmed the purity of the atropinesterase preparations, as Glu was shown to be the only N-terminal residue in the protein sample, both after reaction of the enzyme with dansylchloride by the method of Gros (58) and after automatic degradation of the enzyme in a sequencer. Further confirmation was found in the observation that in each of subsequent degradation steps performed in this sequencer, again only one amino acid was released from the N-terminus.

From the C-terminus only Lys was liberated after incubation of the protein with CPB, whereas CPA+CPB released Phe and Leu as well, in quantities which were almost equimolar to the amount of Lys (table IV.2).

Both N- and C-terminal endgroup determinations corroborate the homogeneity of the protein material.

### IV.2.1.1 Experimental details

PAGE (III.1.2.1), IEF (III.1.2.2) and procedures applied to determine N- and C-terminal endgroups (III.1.5) were performed as described in the indicated sections of chapter III. CPA and CPB hydrolyses were performed on ca. 20 nmol protein. Dansylation of the protein was performed in the presence of 1% SDS as described under IV.6.1.

Table IV.2 Amino acid residues released from atropinesterase during an incubation with CPB and CPA+CPB.

| Enzyme | Incubation time | Amino acids released; nature and relative amounts* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (h) | Met | Val | Ala | Asp | Asn/Gln ** | Glu | Phe | Leu | Lys |
| CPB | 4 | | | | | | | | | 1.0 |
| CPA+CPB | ½ | | | | | | | 0.2 | 0.4 | 0.6 |
| " | 18 | 0.8 | 0.9 | 1.3 | 1.2 | 0.8 | 0.9 | 0.9 | 1.8 | 2.2 |
| " | 18 | 0.9 | 1.1 | 1.5 | 1.2 | 0.9 | 1.0 | 1.0 | 1.9 | 2.0 |

* The amino acid residues released were identified in an automated analyser. The amounts are given relative to the amount of Lys released with CPB.

** A peak was found at the position of Ser, Asn and Gln, which elute close together. The assignment was made because of the high ratio A440/A570.

## IV.3 THE MOLECULAR WEIGHT OF ATROPINESTERASE

To determine the molecular weight of atropinesterase, the experiments performed on SDS-polyacrylamide gels (IV.3.1a) could be used. Before electrophoresis, atropinesterase and reference proteins were preincubated in urea and SDS with or without 2-mercaptoethanol. The relative mobilities (Rf-values) of the reference proteins found in these experiments were plotted against the logarithm of their molecular weights. From the resulting straight line a value of 31,400 ± 2,400 (SD) (10 experiments) could be read for the molecular weight of atropinesterase. The results of one experiment are shown in fig.IV.3. This value agreed very well with the data obtained by Van der Drift in our laboratory in equilibrium sedimentation studies with the analytical ultracentrifuge (42). He determined a molecular weight for the denatured protein of 30,400 ± 1,450 (SEM) in 4.5-8 M urea and of 30,750 ± 1,650 in 3-4 M guanidine.HCl. With the native enzyme, however, he found values of 57,800 ± 1,000, 57,200 ± 2,400 and 61,600 ± 4,250 depending on the calculation method used.

These results indicated that native atropinesterase exists as a dimer composed of two subunits, each consisting of one polypeptide chain with a molecular weight of ca. 30,000.

Additional experiments were performed on thin layers of Sephadex G-75 to verify the estimated molecular weights by gelfiltration. Bovine

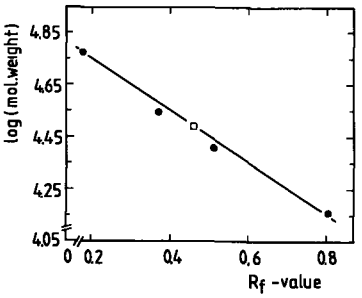



Fig.IV.3. Determination of the molecular weight of atropinesterase by polyacrylamide gelelectrophoresis. A calibration graph was made with the Rf-values found for reference proteins (●) lysozyme, chymotrypsin, pepsin and catalase, from which the molecular weight of atropinesterase (□) was read.

serumalbumin (BSA), chymotrypsinogen, pyruvate kinase and pancreatic RNAse were used as marker proteins. From the plot of log molecular weight versus the mobility relative to BSA, a molecular weight value of 50,000 was estimated. This value is rather low compared to the ultracentrifugation data obtained under non-denaturating conditions (about 60,000), but comes closer to this value than the molecular weight of 39,000 found by Stevens (126) after column chromatography of the native enzyme on Sephadex G-100 with hexokinase and LDH as markers. Chromatography on thin layers of Sephadex G-75 sf in urea, SDS or ammonium sulfate, performed to confirm the molecular weight of the monomer, was not successful since no reproducible values could be obtained.

The molecular weight of the monomer, determined by SDS-PAGE and by use of the ultracentrifuge, was supported by extensive hydrolysis of the protein with CPA+CPB. These exopeptidases released several amino acids from the C-terminus in a ratio of one or two residues per ca. 27,000 molecular weight. The result is summarized in table IV.2, where the quantity released per amino acid is given relative to the amount of Lys released by CPB. This result, together with the fact that in each sequencer step only one amino acid was released and the observation that only one molecular weight species was found in denatured atropinesterase in the ultracentrifuge and by SDS-gelelectrophoresis, strongly suggested that the two monomeric strands in the native enzyme are identical.

It follows from the determination of the catalytic center activity that 1 nmol of active centers corresponds with 30 units of esterase activity. From the amino acid analyses of the whole protein relative

to the amount of enzyme expressed in activity units, a molecular weight per catalytic center of 27,000-29,000 was calculated. This pointed to the presence of one catalytic center per monomer and once more indicated that the two subunits in atropinesterase are identical. Summarizing: the combined results demonstrate that atropinesterase is a dimer with a molecular weight of ca. 60,000. The dimer comprises two identical polypeptide chains, each with a molecular weight of ca. 30,000 and one catalytic center, i.e. one active serine-residue.

IV.3.1 Experimental details

SDS-PAGE of atropinesterase and reference proteins was performed as described under methods (III.1.2.1). Samples were run simultaneously on separate gels under identical conditions. After electrophoresis the length of the gel and the distance moved by the dye were measured before staining, while after destaining the length of the gel and the positions of the blue protein zones were recorded. The Rf-value of each protein was calculated according to Weber et al. (134) by dividing the migration distance of the protein by that of Bromophenol Blue multiplied by the ratio of the gel length before and after staining. The known molecular weights were plotted versus the Rf-values on a semi-logarithmic scale.

Thin layer chromatography on Sephadex G-75 sf was performed on 0.8 mm gels, equilibrated for 16 h with 0.9% NaCl and developed with the same solvent for 6 h at roomtemperature in an LKB apparatus. Proteins were detected on a replica on Whatman 3MM paper with a solution of 1% (w/v) Bromophenol Blue in methanol:acetic acid (9:1, v/v). After staining, the paper was rinsed with 5% acetic acid until a white background was obtained. The relative mobility of each protein was calculated by dividing the migrated distance by the displacement of BSA.

The calibration curves from which the molecular weight of atropinesterase was read, were obtained with at least 4 of the following reference proteins: insulin (5,700), lysozyme (14,300), chymotrypsinogen A (25,740), subtilisin (27,600) pepsin (35,000), ovalbumin (43,000), catalase (60,000), BSA (68,000), pyruvate kinase (57,000), pancreatic RNAse (13,700). (Figures in parentheses refer to the molecular weight as given by Weber and Osborn (134) or Dunker and Ruekert (44)).

## IV.4 THE AMINO ACID COMPOSITION OF ATROPINESTERASE

The amino acid composition of atropinesterase was determined after acidic hydrolysis, for various lengths of time, of different (almost) pure preparations. Cysteine was determined as CMCys after hydrolysis of a protein sample which had been treated with iodoacetic acid. The result given in table IV.3 is the result of numerous determinations. The number of residues of each amino acid was calculated on the

Table IV.3 Determination of the amino acid composition of atropinesterase.

| Amino acid | residues per molecule a) | nearest integer |
|---|---|---|
| Asx | 30.68 | (31) |
| Thr b) | 14.34 | (14) |
| Ser b) | 20.62 | (21) |
| Glx | 19.52 | (19-20) |
| Pro | 15.01 | (15) |
| Gly | 17.19 | (17) |
| Ala | 25.37 | (25) |
| Val c) | 16.13 | (16) |
| Met | 8.76 | ( 9) |
| Ile c) | 14.88 | (15) |
| Leu | 25.30 | (25) |
| Tyr | 7.90 | ( 8) |
| Phe | 11.14 | (11) |
| His | 7.95 | ( 8) |
| Lys | 20.83 | (21) |
| Arg | 7.17 | ( 7) |
| Trp d) | 7.35 | ( 8) |
| Cys e) | 1.62 | ( 2) |
| No.of residues | | 272-273 |

Amino acid analysis was performed on samples of atropinesterase that had been subjected to hydrolysis in 6 N HCl, 110°C, for 24, 48 and 72 h. Unless otherwise indicated, for each amino acid the 3 values were averaged.

a) Values are expressed as residues per molecule assuming a molecular weight of 30,000. Data obtained from 15 amino acid analyses were averaged.
b) Values of Thr and Ser were found by extrapolation to zero time.
c) Ile and Val were taken from the 48 and 72 h hydrolysate, only when values differed from the 24 h hydrolysate.
d) Trp is the mean of 4 analyses, uncorrected for possible destruction (at least 10%).
e) Cys was determined as CMCys and is the mean of 6 analyses, uncorrected for possible destruction (at least 10%).

assumption that the monomeric polypeptide has a molecular weight of 30,000. Then, according to the amino acid analyses the monomer consists of 272-273 residues.

The outcome of the amino acid analyses is consistent with the conclusion reached on the basis of the observations described in the previous section, viz. that the atropinesterase preparations purified according to the procedures used were sufficiently pure to allow for the intended elucidation of the primary structure of the protein, i.e. an amino acid sequence of ca. 270 residues.

IV.4.1 Experimental details

Protein samples (up to 5 nmoles) were hydrolyzed with 6 N HCl for 24, 48 and 72 h at 110°C in the presence of 2% TGA to prevent destruction

of Trp and Met. Incubation periods longer than 24 hours were applied for complete hydrolysis of some peptide bonds, like the ones involving Ile, Leu or Val. Values for Thr and Ser corrected for destruction during hydrolysis were obtained by extrapolation to zero time. Cysteine was determined as CMCys in carboxymethylated atropinesterase (III.1.4.1). Modification of atropinesterase with iodoacetic acid was performed as described in III.1.3.2a.

## IV.5 THE C-TERMINAL SEQUENCE OF ATROPINESTERASE

Residues at the C-terminal side were identified by incubation of the protein with CPB and CPA+CPB. Lys was shown to occupy the C-terminal position, as this was the only residue released with CPB. After incubation with CPA+CPB for a short time, Lys was shown to be preceded by Phe and Leu, whereas extensive hydrolysis with CPA+CPB revealed the identity of 8 residues preceding Phe, Leu and Lys as given in table IV.2. From the results the following tentative C-terminal sequence was deduced: (Met, Val, Ala, Asp, Asn/Gln, Glu, Leu, Lys) Phe-Leu-Lys-OH.

## IV.6 THE N-TERMINAL SEQUENCE OF ATROPINESTERASE

Sequencing of the complete protein from the N-terminus was attempted by application of the stepwise Edman-degradation, both in the manual dansyl-Edman version and in the automated manner, in a commercial sequencer. The automated procedure could be continued much further than the manual method. It elucidated the identity of 29 residues at the N-terminal side:

1 10 17

Glu-Ile-Ile-Pro-Val-Pro-Asp-Glx-Ala-Ala-Trp-Asn-Ala- X -Lys-Lys- X -

18 27

Ile-Glx-Ile-Asn-Asp-Ala-Ile-Lys-Met-Arg-Tyr-Val-Glu-Trp

At the positions 14 and 17 no residues were found; this might point at Ser which is easily destroyed in the identification procedures. Furthermore, at the positions 8 and 19 some ambiguity was encountered, since in both cases Gln as well as Glu was identified. Although the quantity of Gln was small, relative to Glu, its presence strongly suggested that the original residues occupying these positions possessed the amide character. The appearance of both the amide and the acid

at the same position could be caused by the conversion of Gln into Glu during purification of the protein, the same process that was thought to be responsible for the occurrence of more than one protein zone after IEF (IV.3.1b).

All residues were definitely characterized by HPLC studies of the PTH-derivatives and, in 7 cases, by identification of the regenerated amino acids. The sequence of the first 6 residues was confirmed by the manual dansyl-Edman procedure.

IV.6.1 Experimental details

The manual dansyl-Edman degradation of atropinesterase and the automated N-terminal analysis were both carried out with 100 nmol of protein. The sequencer analysis was performed as described in III.1.10.3a. The residues were identified by HPLC; for 27 of the first 31 positions a clear peak of the residue concerned was found, while at positions 8 and 19 a large peak of Glu and a small peak of Gln were observed. For the positions 14 and 17 no peak was found. From the decrease in the amount of Val between positions 5 and 29, a repetitive yield of 94% could be calculated. It has to be noted that from the 6th residue onward, a constant background of some amino acids was found; nevertheless, for the 29 positions mentioned in IV.6 the amino acid concerned could be determined unambiguously. Identification of Lys as the 15th, 16th and 25th residue, Ile at positions 18, 20 and 24 and Arg at 27 was confirmed after conversion of the PTH-derivative into the parent amino acid with HI (III.1.10.3b). Quantitative results are given in table IV.4

Table IV.4 N-terminal sequence of atropinesterase determined with an automated sequencer.

| Cycle | Residue | nmoles | Cycle | Residue | nmoles | Cycle | Residue | nmoles |
|---|---|---|---|---|---|---|---|---|
| 1 | Glu | 70 | 11 | Trp | n.d. | 21 | Asn | n.d. |
| 2 | Ile | 81 | 12 | Asn | 70 | 22 | Asp | " |
| 3 | Ile | 65 | 13 | Ala | 69 | 23 | Ala | 43 |
| 4 | Pro | 81 | 14 | - | - | 24 | Ile | 33 (19) |
| 5 | Val | 80 | 15 | Lys | n.d.(19) | 25 | Lys | n.d.(12) |
| 6 | Pro | 46 | 16 | Lys | n.d.(17) | 26 | Met | 24 |
| 7 | Asp | n.d. | 17 | - | - | 27 | Arg | n.d.( 6) |
| 8 | Glu/Gln | 72 | 18 | Ile | 46 (26) | 28 | Tyr | 29 |
| 9 | Ala | 102 | 19 | Glu/Gln | 26 | 29 | Val | 25 |
| 10 | Ala | 98 | 20 | Ile | 21 (24) | 30 | Glu | 26 |
| | | | | | | 31 | Trp | n.d. |

Procedures for sequencer operation and analysis of fractions are given in III.1.10.3 and IV.7.1.
PTH-amino acids obtained at each cycle were identified by HPLC and in some cases as the parent amino acid following back-hydrolysis. The residue assigned at each cycle was the only major new PTH-amino acid identified. The quantities refer to the amounts found by HPLC; values in parentheses indicate the amounts found after back hydrolysis.
n.d.: not determined; -: no assignment possible.

Dansyl-Edman degradation of the protein was performed as described sub III.1.10.1 for peptides, with slight modification. The first degradation was performed in the presence of 0.5% SDS. After each cycle a sample (in 20 µl 0.5 M $NaHCO_3$, pH 9.9) was incubated with dansylchloride (10 µl, 5 mg/ml) in the presence of 1% SDS for 5 h at 37°C. Subsequently, the protein was precipitated by addition of 40 µl 20% trichloroacetic acid and 4 µl BSA (20 mg/ml) as a carrier protein. After standing in ice for 1 h, the precipitate was centrifuged and washed twice with 50 µl 1 N HCl. During the identification of the 4th residue, next to DNS-Pro the dansylated dipeptide Pro-Val was obtained (see also III.1.10.1). The dried pellet was hydrolyzed with 6 N HCl. Identification of amino acids beyond the 6th residue was without success owing to the presence of a background of DNS-amino acids which eliminated the possibility to identify the next residue.

## IV.7 PRESENCE OF SUGAR RESIDUES

To establish whether sugar residues are attached to atropinesterase, the sugar content was determined by the method of Dubois et al. (43) with fructose as the reference compound. As a control, trypsin was used. For both proteins about equal quantities of sugar were measured (< 2%), indicating that atropinesterase does not contain a significant amount of sugars.

## IV.8 CONCLUSIONS

The results of the study described in this chapter confirm the original impression that atropinesterase would be a suitable member of the group of serine-esterases to be studied in a sequence investigation aimed at a structural and evolutionary comparison with the well-studied serine-proteases. Although the native enzyme is a dimer, both subunits are identical, each containing one active serine residue. With regard to size the monomer is comparable to, for instance, subtilisin. Similar to this protease, monomeric atropinesterase consists of one polypeptide chain. Not unexpectedly for a protein of that size, straightforward sequencing of the whole chain was not possible, so selective fragmentation by enzymes and/or chemical agents had to be applied. According to the amino acid composition, suitable residues for a specific cleavage appeared to be present in numbers allowing a degradation into a limited number of peptides, e.g. bond-scission at the 7 Arg-residues would result in 8 fragments with an average length of 30-40 residues.

## CHAPTER V THE AMINO ACID SEQUENCE AROUND THE ACTIVE SERINE OF ATROPINESTERASE

Isolation and analysis of a cyanogen bromide peptide and a peptide obtained after digestion of the maleylated protein with trypsin.

As a first approach to the structural study of the active center of atropinesterase, a.o. to see whether sequence homology with other serine-hydrolases existed, the isolation and sequencing of fragments containing the active serine residue was attempted. In earlier attempts in our laboratory, the enzyme had been radioactively marked at this serine with the $^{32}P$-labeled inhibitor Soman. Subsequent breakdown with various proteolytic enzymes was tried in those experiments, but a suitable $^{32}P$-peptide could not be isolated due to continuing splitting up of peptides and release of the $^{32}P$-substituent. Partial hydrolysis with HCl then led to a few, more stable di- and tri-peptides from which the sequence His-Ser-Met-Gly was deduced.

In the experiments presented here, the -presumably- more stably bound organophosphate DFP was used to label the active serine with $^{32}P$. The first approach with CNBr-cleavage led to an active serine-containing fragment (CN-a) that could be well recognized and isolated, so that in the final characterization of this dodecapeptide -for practical reasons- labeling was omitted.

The sequence of this rather small fragment did not show obvious homology with the sequence around the active serine of other enzymes; as serine was penultimate, comparison in only one direction was possible, into a region where much variability is allowed in the prothrombin-related proteases and in the subtilisins.

As no conclusion could be drawn from these results, the isolation of a longer fragment was attempted by digestion of the maleylated protein with trypsin. Again labeling with [$^{32}P$]-DFP was used in the initial experiments to recognize the active serine-containing fragment (Tm-a). Owing to the size and further properties of this peptide of about 60 residues, several problems had to be solved before a suitable isolation procedure could be worked out. The second part of this chapter describes this isolation and the subsequent sequencing studies.

## V-I.1 CYANOGEN BROMIDE CLEAVAGE OF ([$^{32}$P]-DFP-LABELED) ATROPINESTERASE; FRACTIONATION OF THE FRAGMENTS; ISOLATION AND CHARACTERIZATION OF A PEPTIDE COMPRISING THE ACTIVE SERINE RESIDUE

### V-I.1.1 THE ACTIVE SERINE-CONTAINING PEPTIDE OBTAINED AFTER DEGRADATION OF [$^{32}$P]-DFP-INHIBITED ATROPINESTERASE

In the first experiments, atropinesterase inhibited with [$^{32}$P]-DFP was used for the degradation by CNBr. The radioactive label attached to the active serine made it possible to identify the fragment(s) containing this serine and to follow the peptide(s) through the purification steps. Different portions of labeled atropinesterase, ranging from 0.07 to 0.50 µmol (corresponding to 2-15 mg) were cleaved with CNBr. The resulting CN-peptides were fractionated on a column of Sephadex G-50 sf. In fig.V.1A the reproducibly obtained elution patterns with regard to the A280, the ninhydrin colour measured in the peptide analyzer (A570) and the $^{32}$P-content of the fractions (two experiments), have been combined. Three $^{32}$P-peaks were obtained. The main variation was observed in peak III, which could be ascribed to the hydrolysis product of DFP, viz. diisopropyl phosphate (DIP), which always was present in the digests of DFP-inhibited atropinesterase. The major part of the peptide-bound radioactivity was recovered in peak II, which evidently represented the active serine peptide. This peak coincided with peak IIa in the elution pattern of polypeptide material (A570). In a representative experiment it contained 79% of the bound $^{32}$P. Peak I, with 19% of the peptide-bound $^{32}$P, probably contained undigested protein or partially degraded fragments.

In one of the experiments shown in fig V.1A (digest B), fractions of peak II were pooled as indicated. The pool was fractionated by paper-electrophoresis at pH 3.6. Five radioactive, ninhydrin-positive spots were found. The eluates contained 62, 8, 23, 7 and 1%, respectively of the $^{32}$P present before electrophoresis. The presence of 2 spots could be explained by the presence of Hse and Hse-lactone at the C-terminus of the peptide, but no explanation could be given for the other spots. The eluate of the strongest spot was refractionated by paper-

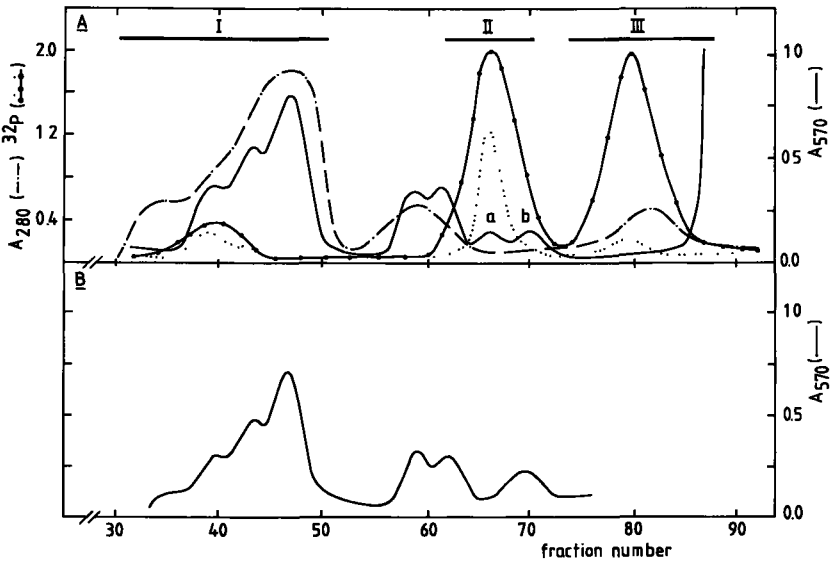



Fig.V.1. Fractionation of CNBr-degraded atropinesterase, [$^{32}$P]-DFP-labeled (panel A; digest A and B) or unlabeled (panel B; digest C), on Sephadex G-50 sf.
The mixture of CN-peptides from 70 nmol (digest A) and 0.5 µmol (digest B) of $^{32}$P-labeled atropinesterase, and from 40 nmol of unlabeled protein (digest C) was chromatographed on the same column (160 x 0.9 cm). Elution was with 5% acetic acid. The eluates were monitored for the presence of ninhydrin-positive material in the PA (A570:——; digest A and C), for A280 containing material (-.-.; digest B) and for $^{32}$P-content (digest A:...; digest B:●—●). The A570 is given in recorder-units which are only roughly comparable to absorbance. $^{32}$P (...) is given in arbitrary units; $^{32}$P (●—●) is in 10 cpm/ 50 µl. Fractions of digest B in peak IIa were pooled according to the horizontal bar marked II. For details, see V-I.1.2.1).

electrophoresis at pH 3.6. Again more than 2 radioactive spots were found. This phenomenon could be repeated several times. Purification of the strongest spot by paperelectrophoresis at pH 6.5 yielded the active serine-containing peptide in an almost pure state according to endgroup determination with dansylchloride. This showed DNS-Gly as the only DNS-amino acid, demonstrating Gly to be the N-terminal residue. The amino acid composition confirmed the homogeneity of the peptide (table V.1). Per mol peptide 1.1 mol $^{32}$P was found and 1.0 mol Ser (after correction for losses during hydrolysis) suggesting that this amino acid is the active serine residue present in atropinesterase.
For sequence determinations, larger amounts of peptide material were required. In view of problems encountered during the purification of

Table V.1 Amino acid composition of CN-a, the active serine peptide, isolated after CNBr-cleavage of [$^{32}$P]-DFP-labeled and unlabeled atropinesterase.

| Amino acid | [$^{32}$P]-CN-a[a)] | CN-a[a)] | CN-a[b)] | c) |
|---|---|---|---|---|
| Asx | 1.37 | 1.09 | | (1) |
| Thr | 2.07 | 2.03 | 2.0 | (2) |
| Ser | 0.95 | 1.06 | 1.7 | (1) |
| Glx | | | | |
| Pro | | | | |
| Gly | 2.02 | 2.16 | 2.1 | (2) |
| Ala | | | | |
| Val | 0.91 | 1.08 | 1.1 | (1) |
| Met | 0.98 | 0.99 | 1.0 | (1) |
| Ile | 0.84 | 0.74 | 1.0 | (1) |
| Leu | 1.10 | 0.99 | 1.2 | (1) |
| Tyr | | | | |
| Phe | | | | |
| His | 1.76 | 1.85 | 1.7 | (2) |
| Lys | | | | |
| Arg | | | | |
| Trp | | | | |
| $^{32}$P | 1.11 | | | |
| No.of residues | 12 | 12 | | 12 |
| Yield | 6% | 18% | | |
| $NH_2$-terminus | Gly | Gly | | |

N-terminal residues were determined by the dansyl procedure.
Yields refer to degraded amount of atropinesterase.

a) Residues were determined after hydrolysis of ca. 20 nmol in 6 N HCl for 24 h at 110°C and are reported as mol per mol peptide (see III.1.4.2). Ser and Thr were not corrected for degradation losses. Met was determined as Hse. Amino acids present at a level less than 0.2 mol/mol peptide are not shown.

b) Residues obtained after exhaustive degradation with LAP. In this case Asn may be formed which is determined together with Ser.

c) Figures in parentheses show the number of residues of each amino acid found by sequence studies.

the $^{32}$P-labeled fragments and because of the practical limitations involved in the handling of radioactive preparations, the subsequent experiments with CNBr were performed with unlabeled, uninhibited atropinesterase.

V-I.1.1.1 Experimental details

The amount of [$^{32}$P]-DIP was determined either by paperchromatography (see III.1.1.3) or by paperelectrophoresis at pH 3.6 (45 min at 2000 V). For more details see V-I.1.2.1.

V-I.1.2 THE ACTIVE SERINE-CONTAINING PEPTIDE OBTAINED AFTER CNBr-CLEAVAGE OF ATROPINESTERASE

In a typical experiment 2.1 µmol of atropinesterase (63 mg) was cleaved with CNBr. Two percent of the digest was chromatographed on the same column of Sephadex that had been used for the digests of [$^{32}$P]-DFP-inhibited atropinesterase. The resulting elution profile of ninhydrin-positive material (fig.V.1B) strongly resembled the earlier obtained pattern (fig.V.1A) but for the position of the active serine peptide. No peak was observed at the position of the $^{32}$P-peptide (peak IIa); the unlabeled fragment appeared to be eluted later, presumably because of its smaller size (mol.weight is 164 lower), and (almost) coincided with peak IIb which was correspondingly enlarged.

The main portion of the digest (98%) was fractionated on a larger column which gave essentially the same result (fig.V.2). Fractions expected to contain the active serine peptide were pooled as indicated, concentrated by lyophilization and fractionated on Sephadex G-25 sf as

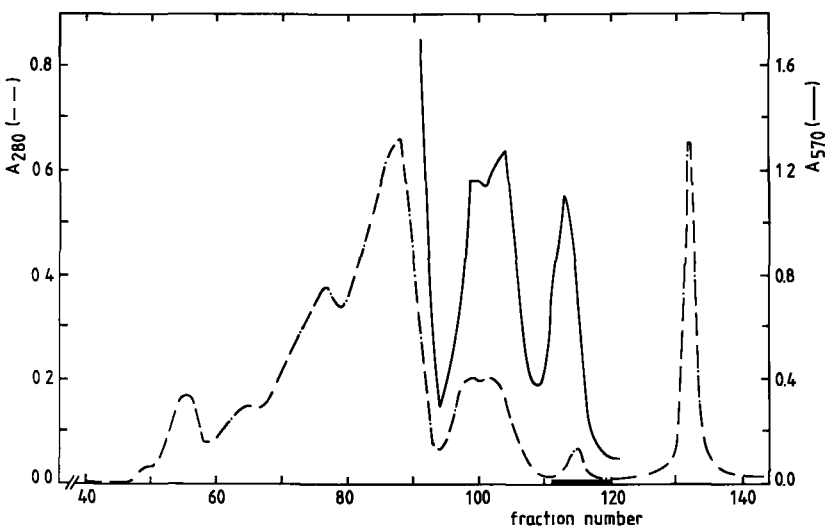



Fig.V.2. Fractionation of CN-peptides of atropinesterase on Sephadex G-50 sf. CNBr-degraded atropinesterase (2.06 µmol; digest C of fig.V.1) was chromatographed on a Sephadex G-50 sf column (110 x 2.3 cm) by elution with 5% acetic acid. Fractions (3.3 ml) were monitored at 280 nm (-·-·-) and fractions 87-121 also for the presence of ninhydrin-positive material, in the PA (A570:——).

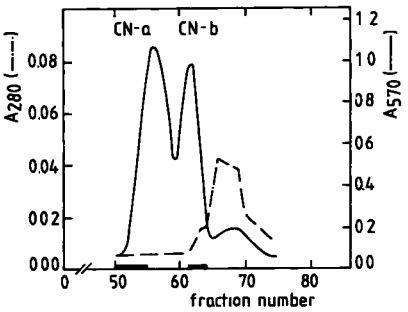



Fig.V.3. Purification of CN-a, the CN-peptide containing the active serine. The material in the pooled fractions of fig.V.2 were fractionated on a column (280 x 1.5 cm) of Sephadex G-25 sf, eluted with 1% acetic acid. Peptides were detected on line by reaction with ninhydrin in the PA. Fractions of 3.4 ml were collected which were monitored at 280 nm (-•-•-). The fractions indicated by the horizontal bars contained pure CN-a and pure CN-b, resp.

shown in fig.V.3. Two major, well-resolved ninhydrin-positive peaks were obtained. The leading half of the first peak (a) was pooled; it contained a pure peptide (CN-a) with the same amino acid composition as the active serine fragment isolated from [$^{32}$P]-DFP-inhibited atropinesterase (see table V.1). Also the tailing half of peak b was pooled; it contained a pure decapeptide, CN-b, with a quite different amino acid composition (the further characterization of this peptide will be discussed in chapters VI and VII). In the fractions between the two pools mixtures of both peptides were found. In the material used for the separation on Sephadex G-25 sf, the yields of CN-a and CN-b were 43 and 76%, respectively, with respect to the degraded amount of atropinesterase. Estimation of these yields was possible because each peptide contained amino acid residues not present in the other, while no other peptides were present in substantial yield. In view of the about 80% $^{32}$P consistently found in peak II with the labeled enzyme, a higher yield of CN-a was expected; possibly, fractions containing CN-a have been omitted during pooling, but it appears unlikely that this could be the only explanation. In pool a, comprising less than half of peak a, 0.37 µmol of CN-a was recovered, i.e. 18% overall yield; no corrections were made for losses during the isolation procedure.

In order to confirm the amino acid composition determined for CN-a and to investigate the possibility that the Asp residue found after acid hydrolysis occurred in the peptide as Asn, a degradation with LAP was performed. All constituting amino acids were released in stoichiometric amounts (table V.1), but instead of Asp an increased quantity of

Ser was found upon amino acid analysis. As Asn co-eluted with Ser, and because of the high A440/A570 ratio observed for the Ser peak (3) the presence of Asn became apparent.

V-I.1.2.1 Experimental details

Inhibition of atropinesterase with [$^{32}$P]-DFP and cleavage of labeled and unlabeled atropinesterase with CNBr was performed as described in chapter III.

Fractionation of the CN-peptides of [$^{32}$P]-DFP-labeled atropinesterase (fig.V.1A; digest A and B)) and of 2% of the unlabeled protein (fig.V.1B) was carried out on a column (160 x 0.9 cm) of Sephadex G-50 sf eluted with 5% acetic acid. The labeled CN-peptides were eluted at a flow rate of 5 ml/h. Of the effluent of digest A, 30% was used for automatic analysis (on line) by reaction with ninhydrin following alkaline hydrolysis, in the peptide analyzer (PA) (fig.V.1A; A570). In this experiment, also the radioactivity was monitored continuously at the outlet of the column with a Geiger-Müller counting tube. The $^{32}$P elution pattern of digest B was obtained by liquid scintillation counting (LSC) of aliquots (4.5%) of each fraction; fraction size was 1.1 ml. With this digest, also the A280 of the fractions was measured. Fractions containing the active serine peptide, peak II, were pooled, as indicated by the bar. CN-peptides of the unlabeled protein (digest C, fig.V.1B) were eluted at a flow rate of 8.3 ml/h; 22% of the effluent was used for on line detection of peptides (A570).

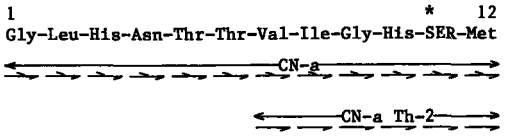
The major portion of the peptide mixture of the unlabeled protein (98%; 2.06 µmol; 62 mg) was chromatographed on a wider column (110 x 2.3 cm) of Sephadex G-50 sf, eluted with 5% acetic acid at 33 ml/h. Fractions of 3.3 ml were collected; their A280 was measured and aliquots (1.1%) of fractions 87-121 were analyzed by reaction with ninhydrin in the PA (sample uptake 1.6 ml/h). Results are given in fig.V.2. Fractions expected to contain CN-a were pooled as indicated, concentrated and fractionated on a column (280 x 1.5 cm) of Sephadex G-25 sf in 1% acetic acid at a flow rate of 22 ml/h (see fig.V.3). Seven percent of the effluent was used for colour reaction with ninhydrin in the PA. The remainder of the eluate was collected in fractions of 3.4 ml of which the A280 was measured. Two pools were made, as indicated by the bars; these were shown to contain pure CN-a and pure CN-b, respectively.

LAP hydrolysis of CN-a was as described in III.1.5.1a; digestion of 17 nmol was for 6 h at 37°C.

V-I.2 THE AMINO SEQUENCE OF THE ACTIVE SERINE PEPTIDE CN-a

Dansyl-Edman degradation of CN-a identified the N-terminal sequence of 9 residues and the residue at position 11 (fig.V.4). According to LAP results (V-I.1.2.1), DNS-Asp found for position 4 should originate from Asn. As CN-a was formed by CNBr-cleavage, the Hse-residue present should be C-terminal; then, according to the amino acid composition

Fig.V.4 The amino acid sequence of peptide CN-a.




The result was obtained by sequence studies on CN-a and on the fragment CN-a Th-2 obtained after digestion with thermolysin.

—→ Residue identified as DNS-amino acid.
⇝ Amino acids released by leucinaminopeptidase (LAP).
* The active serine residue.

the 10th residue must be His. To verify this conclusion, the isolation of a C-terminal fragment of CN-a was attempted, after digestion with thermolysin. Fractionation of the peptide mixture by paperelectrophoresis at pH 3.6 resulted in the isolation of the undigested peptide CN-a (yield 13%), a mixture of CN-a and the N-terminal thermolytic peptide, CN-a Th-1 (yields 38 and 44%, respectively), which were not further characterized, and the pure C-terminal peptide CN-a Th-2, the most basic spot (yield 27%). For CN-a Th-2 the following composition was found: Ser(0.94) Gly(1.19) Val(1.05) Met(1.08) Ile(1.01) His(0.91). Successful dansyl-Edman degradation of CN-a Th-2 completed the sequence of the active serine peptide as indicated in fig.V.4. The result is in agreement with the sequence His-Ser-Met-Gly obtained earlier in our institute after partial acid hydrolysis (see introduction chapter V).

V-I.2.1 Experimental details

Thermolytic digestion of CN-a was carried out in 0.02 M Tris.HCl, pH 8.0, containing 1 mM $CaCl_2$, at a peptide concentration of 1 µmol/ml (1.2 mg/ml) for 2h at 37°C, with 0.04 mg thermolysin per µmol peptide. The reaction was stopped by lyophilization. Details are summarized in table III.1.

Fractionation of CN-a Th-peptides was carried out by high voltage paperelectrophoresis (HVE) at pH 3.6 for 90 min at 4000 V; peptides were detected on guidestrips with fluorescamine and ninhydrin (II) as described in III.1.9.2.

The amino acid composition of CN-a Th-2 was determined after hydrolysis of ca. 10 nmol for 18 or 48 h; given values are the mean of 2 18-h hydrolysates. Values for Val and Ile were taken from the 48 h hydrolysate.

All procedures generally used in the sequencing of peptides like the dansyl-Edman procedure, performance of HVE, preparation of reagents used for detection of peptides, etc. were as described in chapter III.

## V-II.1 TRYPTIC DIGESTION OF MALEYLATED ([$^{32}$P]-DFP-LABELED) ATROPINESTERASE: FRACTIONATION OF THE FRAGMENTS AND PURIFICATION OF THE PEPTIDE CONTAINING THE ACTIVE SERINE

In the hope to obtain a larger active serine-containing fragment, digestion with trypsin of maleylated atropinesterase was attempted. This approach was successful in that a peptide of about 60 residues was formed. The isolation of this fragment, however, proved to be rather problematic and the procedure described here has been preceded by numerous unsuccessful attempts. In the presentation of the purification of Tm-a, occasionally earlier results are briefly discussed to indicate why certain methods were chosen.

### V-II.1.1 LOCALIZATION OF THE [$^{32}$P]-DFP-LABELED ACTIVE SERINE PEPTIDE AFTER GELFILTRATION OF THE TRYPTIC DIGEST

In a pilot experiment atropinesterase was inhibited with [$^{32}$P]-DFP prior to maleylation and digestion with trypsin, to make the localization possible of the tryptic peptide that contains the active serine. In the experiment shown, 80 nmol $^{32}$P-labeled atropinesterase (2.4 mg) was maleylated with 1,4-[$^{14}$C]-maleic anhydride. [$^{32}$P]-DIP and 1,4-[$^{14}$C]--maleic acid were removed by gelfiltration on Sephadex G-50 sf. The resulting material was lyophilized, dissolved in Tris.HCl buffer and denatured in boiling water. Portions were digested with trypsin. The tryptic fragments (Tm-peptides) were separated on a Sephadex G-50 sf column. A representative elution pattern of $^{32}$P and peptide material is shown in fig.V.5. Five $^{32}$P-peaks (I through V) were found, containing 7, 55, 20, 5 and 13% of the eluted $^{32}$P-radioactivity, respectively. Peak II, the major radioactive peak, was assumed to contain Tm-a, the tryptic peptide with the active serine residue. According to the elution volume, top I probably contained undigested protein. From the elution volume of peak II a molecular weight of about 7,000 was roughly estimated. (This estimate could be slightly high due to the presence of negatively charged maleyl groups; maleylated peptides usually are eluted earlier from Sephadex columns than "normal" peptides of the same size). The $^{32}$P in peak V was identified as DIP.

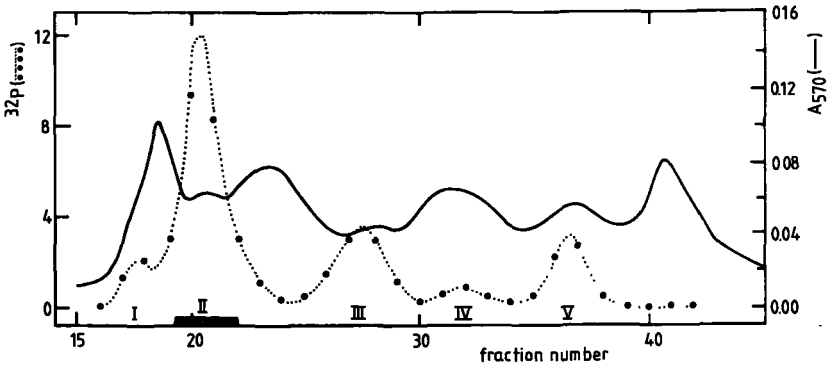



Fig.V.5. Fractionation of Tm-peptides of 20 nmol [$^{32}$P]-DFP-labeled atropinesterase on Sephadex G-50 sf. The column was eluted with 0.01 M Tris.HCl, pH 8.0. Fractions of 1.8 ml were collected. The effluent was analyzed for ninhydrin-positive material in the PA (A570:——) and for radioactivity ($^{32}$P:...). Aliquots (17%) were used for quantitative determination of the radioactivity by LSC; $^{32}$P (●.●) is in $10^3$ cpm/0.4 ml.

[The appearance of the smaller radioactive fragments could be due to aspecific splitting of one or more peptide bonds in Tm-a, as was suggested by results of subsequent experiments. When conditions were kept constant, small variations were discovered in the distribution of the $^{32}$P; sometimes the second $^{32}$P-peak was slightly reduced, while the third peak was correspondingly enlarged. When tryptic digestion was continued for more than 6 h, however, the $^{32}$P-radioactivity shifted to peaks with larger elution volumes. Suprisingly, the ninhydrin profile and $^{14}$C-pattern (not shown) underwent only minor variation. The constancy of the ninhydrin and $^{14}$C-profile in preparations showing relative large variations in the $^{32}$P-pattern suggested a preferential rupture of some peptide bonds in [$^{32}$P]-Tm-a.]

For further purification and characterization of intact Tm-a, larger amounts of material were needed. Therefore, the procedures were repeated with greater quantities (ca. 3 µmol) of atropinesterase. For the same reasons as were mentioned in V-I.1.1, in subsequent experiments 1,4-[$^{14}$C]-maleylated atropinesterase was digested with trypsin without the preceding modification with [$^{32}$P]-DFP. (Omission of the $^{32}$P-label also facilitated quantitative determination of the $^{14}$C present in the protein).

V-II.1.1.1 Experimental details

Inhibition of atropinesterase with [$^{32}$P]-DFP and maleylation were performed as described in chapter III. Before maleylation the preparation still contained 31% of the $^{32}$P as DIP. 1,4-[$^{14}$C]-maleic acid and [$^{32}$P]-DIP were removed on a column (150 x 0.9 cm) of Sephadex G-50 sf with 0.01 M Tris.HCl, pH 8.0, as the eluent. Fractions of 2.1 ml were collected and their A280 and A250 were measured . Aliquots were used to determine the $^{32}$P- and $^{14}$C-radioactivity. Of the eluted $^{14}$C, 95% was present in the peak of maleic acid; 38% of the eluted $^{32}$P was [$^{32}$P]-DIP. The 31% $^{32}$P-[DIP] found by paperchromatography before maleylation was in good agreement with the 38% found after chromatography of the maleylated protein on sephadex G-50 sf. The lyophilized pool containing 60 nmol 1,4-[$^{14}$C]-maleylated [$^{32}$P]-DFP-labeled protein, was dissolved in 0.6 ml 0.03 M Tris.HCl, pH 8.0, and denatured (see III.1.3.1a). Portions of the denatured protein were digested with trypsin as described in III.1.6.3. The resulting Tm-peptides were fractionated on a column (150 x 0.9 cm) of Sephadex G-50 sf in 0.01 M Tris.HCl, pH 8.0. Elution was with the same buffer at 8.1 ml/h. Part of the effluent (22%) was sacrificed for on line detection of ninhydrin-positive material in the PA. The remaining 78% were collected in fractions of 1.8 ml. $^{32}$P-radioactivity was monitored continuously at the outlet of the column with a Geiger-Müller counting tube. $^{14}$C and $^{32}$P were quantitatively measured by LSC of an aliquot (17%) of each fraction.

V-II.1.2 ISOLATION AND CHARACTERIZATION OF Tm-a

In a representative experiment 2.5 µmol atropinesterase (75 mg) was modified with 1,4-[$^{14}$C]-maleic anhydride (the radioactive label facilitated the check on the reproducibility of the tryptic digestion from the $^{14}$C-elution pattern of Sephadex columns). Excess maleic acid was removed on Sephadex G-50 sf. Protein-containing fractions were pooled, concentrated by lyophilization, denatured by heating and digested with trypsin.

[As in this type of experiment the $^{14}$C-content, the absorbance measurements at 280 and 250 nm (51) and the absence of ε-dansyl-Lys after dansylation of the maleylated protein usually indicated that all Lys-residues had been protected, exclusive cleavage at the Arg-bonds was expected. The extent to which this occurred, and the specificity of the tryptic action were checked by determination of newly formed C-termini with CPB and CPA, respectively. In every experiment with trypsin all Arg-bonds appeared to be ruptured, since all Arg-residues present in the protein had become C-terminal. However, in the early experiments breakage at other bonds also

occurred extensively, as next to Arg many other amino acids had become C-terminal. By changing experimental conditions and the use of different trypsin preparations, the specificity could be improved considerably, to the extent that cleavage was restricted to Arg-bonds, as CPA released less than one molecule of any other amino acid was per molecule of protein, as was the case in the experiment presented here.]

As a first fractionation step, half of the tryptic digest was chromatographed on a Sephadex G-50 sf column. A typical elution profile is shown in fig.V.6. The partition coefficients (Kav-values) of the fractions of peak II in this figure, ranging from 0.11-0.21, were equal to those of the major $^{32}$P-top, peak II in fig.V.5. When the $^{32}$P-radioactivity in this peak II was correctly attributed to the presence of $^{32}$P-labeled Tm-a, peak II in fig.V.6 should comprise the unlabeled active serine peptide, since in this case the absence of the diisopropyl phosphorylgroup was not expected to noticeably affect the over-all size of this peptide (difference of 164 on a total mol.weight of about 7,000).

The pooled central fractions of the peak (nrs. 85 through 88) were studied with respect to the N-terminal sequence present; by dansyl-Edman degradation Gly-His-Gly was found despite contamination with some less intensive DNS-amino acids, probably due to the presence of a few contaminating tryptic peptides. The results suggested that the central fractions of peak II contained one major product, presumably the active serine peptide Tm-a.

[To ascertain that this major product was correctly identified as Tm-a, an earlier but identically obtained peak II preparation was subjected to CNBr-cleavage after demaleylation, to see whether the active serine CN-peptide, CN-a, would be formed. From the digest indeed a reasonable amount of CN-a could be isolated, which demonstrated unambiguously the correct assignment. The isolation of pure CNBr-fragments from a preparation of Tm-a is described in V-II.2.2.1.]

Subsequently, several fractions between fraction number 70 and 100 were investigated -via the dansyl method- for the N-terminal amino

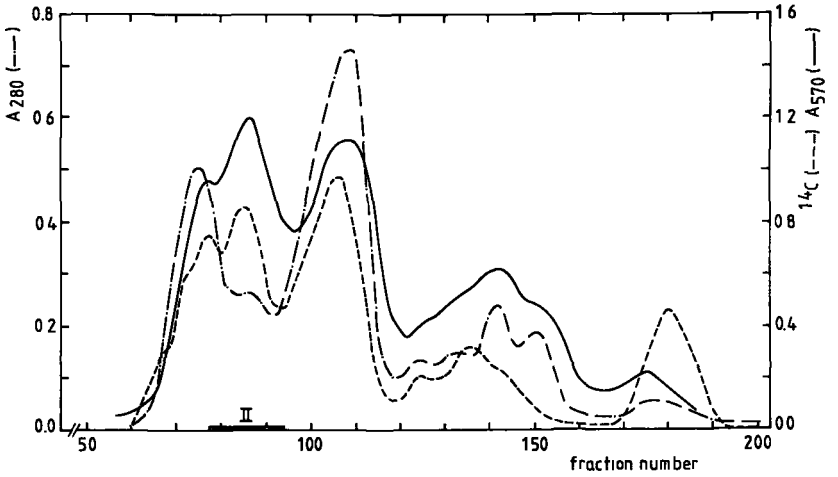


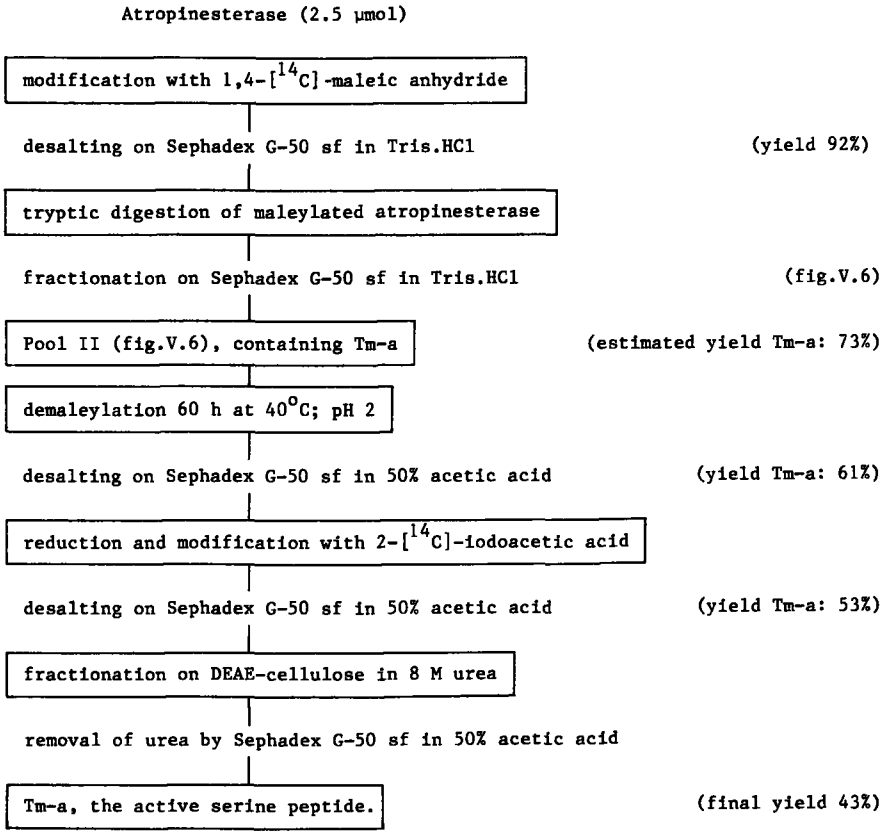

Fig.V.6. Fractionation of Tm-peptides from 1.2 µmol 1,4-[$^{14}$C]-maleylated atropinesterase on Sephadex G-50 sf. The column (136 x 2.3 cm) was eluted with 0.01 M Tris. HCl, pH 8.0. Fractions of 2.4 ml were monitored at 280 nm (-•-•-). Aliquots (1.3%) were analyzed in the PA (A570:——); aliquots (1.0%) were removed for LSC, $^{14}$C (---) is in $10^3$ cpm/50 µl.

acid present; all fractions containing Gly as the predominant N-terminal residue were pooled (fig.V.6; pool II). The chromatography and further procedures were repeated with the second portion of the digest, and the two pools II were combined. This material was subjected to a series of modification and fractionation steps in order to isolate pure Tm-a in a reasonable yield. The whole procedure, which is shown in scheme V, was adopted on the basis of preceding attempts. The selection of some of the steps is discussed in side-remarks to the presentation of this purification study.

Preparatory to the fractionation of pool II, first the peptide material was demaleylated because all earlier attempts to purify maleyl-Tm-a from this pool had been unsuccessful.

[These attempts had failed because of bad separation or poor yields. Nevertheless, fractionation on DEAE-Sephadex once resulted in the isolation of a small amount of almost homogeneous Tm-a, as judged by the presence of only one N-terminal residue (Gly); yield was 18%. The amino acid analysis indicated a length of about 60 residues and the occurrence of the only cystine of atropinesterase.

Scheme V. Isolation of Tm-a, the peptide containing the active serine.




However, the purification method appeared not applicable for the isolation of Tm-a at a scale allowing sequence studies.]

The 1,4-[$^{14}$C]-maleylgroups were removed by prolonged incubation in 33% acetic acid. A turbid solution remained which was clarified by increasing the acetic acid concentration to 50%. $^{14}$C-maleic acid was removed by gelfiltration; peptide-containing fractions were pooled and lyophilized.

The demaleylated material was dissolved in 0.02 M Tris.HCl, pH 8.5, containing 8 M urea. About 95% of the solution was treated with

dithioerythritol (DTE) to reduce the S-S bridge present in Tm-a, and the SH-groups formed were modified by reaction with 2-[$^{14}$C]-iodoacetic acid. The remaining 5.2% of the peptide mixture was treated in an identical manner, except that reduction was omitted; this sample will be discussed later. Excess reagent was removed by gelfiltration on Sephadex G-50 sf in 50% acetic acid. The pooled fractions contained 2.7 µmol CMCys as determined by amino acid analysis and 3.7 µmol [$^{14}$C]-carboxymethylgroups as determined by liquid scintillation counting. The estimated yield of Tm-a in the pool was 61%.

[The fractionation of the demaleylated peptide mixture was greatly hampered by the very poor solubility in suitable aqueous solvents. The material was well soluble in concentrated organic acids, e.g. 50% acetic acid, but purification methods in which media of this kind could be used proved ineffective. Solubilisation in dilute buffers was possible when a high concentration of urea was added, and in these solvents fractionation appeared feasible. However, this addition excluded the use of the PA for detection. Furthermore, the low molar absorbance at 280 nm of Tm-a, caused by the absence of Trp and a low Tyr-content, eliminated the use of A280 for the detection of this peptide among the much stronger absorbing contaminating fragments. A different and even specific detection method seemed accessible by using the presence in Tm-a of the single cystine of atropinesterase, which could be converted into 2 $^{14}$C-labeled CMCys residues.]

The pool with carboxymethylated Tm-a was lyophilized, the residue was taken up into 3 ml 0.01 M Tris, containing 8 M ureum and the pH was readjusted to 8.5 with NaOH. Some material remained undissolved, which could not be removed by centrifugation (15 min at 20,000 x g; flotation occurred, and immediate remixing when the tubes were touched). The solution, together with the turbidity, was applied to a DEAE-cellulose column equilibrated with 8 M urea in Tris.HCl buffer. Elution with the same buffer with an increasing NaCl concentration resulted in the profile as shown in fig.V.7. The $^{14}$C-radioactivity was eluted in one well-separated peak (a) and 3-5 overlapping peaks (b, c, d).

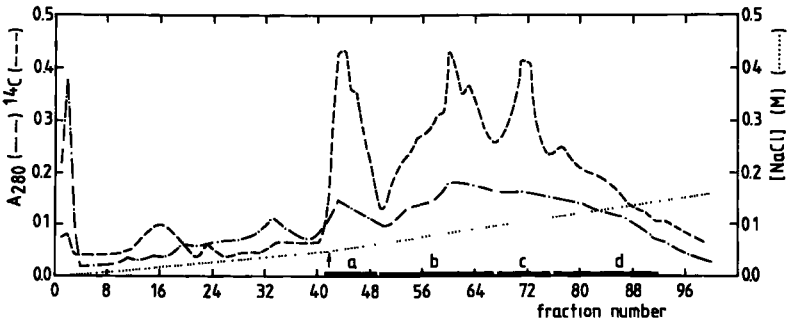



Fig.V.7. Fractionation of demaleylated, reduced and $^{14}$C-carboxymethylated Tm-peptides on DEAE-cellulose. The column (8.8 x 0.9 cm) was eluted with 0.01 M Tris.HCl, pH 8.0, containing 8 M urea and a linearly increasing concentration of NaCl. Fraction size: 1.6 ml. Exact conditions and methods used for peptide detection are given in the text. A280 (-·-·-), $^{14}$C-radioactivity (---) is in $10^3$ cpm/20 µl; calculated NaCl concentration (...).

The material was fractionated as indicated in 4 pools, which contained 16% (a), 33% (b), 18% (c) and 17% (d) of the $^{14}$C brought onto the column. Total yield 84%.

The pools were freed of urea and salts by gelfiltration. After dansylation of pool a, only Gly was found as the N-terminal residue, while after incubation with CPA and CPB for 30 min three C-terminal residues Val, Ser and Arg (table V.6) were found in equimolar quantities. Amino acid analysis showed a well-fitting amino acid composition as given in table V.2. On basis of these results it was concluded that peptide Tm-a had been isolated in pure state, in a workable quantity of 0.2 µmol, be it with the low over-all yield of 8%.

According to amino acid analyses, Tm-a was the main component (> ca. 80%) also of the pools b, c and d, with yields of 19, 9 and 7%, respectively, giving a total yield for Tm-a of 43%. The identity of Tm-a in these pools was confirmed in later experiments, on the basis of peptides formed by secondary cleavages. From the same results, and from the amino acid analyses, the presence of at least one contaminating peptide in pools b, c and d was concluded. The contamination seemed to increase with higher fraction number. No satisfactory explanation has been found for the distribution of Tm-a over 4 different peaks. Possibly molecules with increased acidity were present due to incomplete removal of maleylgroups or partial deamidation of amide

Table V.2 Amino acid composition of Tm-a, the active serine peptide, isolated after tryptic digestion of maleylated atropinesterase.

| Amino acid | amount (mol/mol) | number of residues |
|---|---|---|
| Asx | 6.07 | (6) |
| Thr a) | 3.81 | (4) |
| Ser a) | 6.70 | (7) |
| Glx | 1.42 | (1) |
| Pro | 2.21 | (2) |
| Gly | 6.88 | (7) |
| Ala | 3.25 | (3) |
| Val b) | 5.21 | (5) |
| Met c) | 3.09 | (3) |
| Ile b) | 3.78 | (4) |
| Leu | 2.26 | (2) |
| Tyr | 1.95 | (2) |
| Phe | 1.98 | (2) |
| His | 3.79 | (4) |
| Lys | 3.14 | (3) |
| Arg | 0.94 | (1) |
| Cys d) | 1.79 | (2) |
| No. of residues | 58 | 58 |
| No.of $^{14}$C-groups | 2.20 | |
| Yield | 43% | |
| $NH_2$-terminus | Gly | |

Amino acid analyses were performed after 24 h and 48 h hydrolysis. Residues are reported as mol per mol peptide, according to III.1.4.2. Figures in parentheses show the number of residues of each amino acid found by sequence studies. Given values are the mean of 3 amino acid analyses.
a) Values of Ser and Thr were found after extrapolation to zero time.
b) Values for Val and Ile were taken from the 48 h hydrolysate.
c) Met was determined after hydrolysis in the presence of 2% TGA.
d) Cys was determined as CMCys.

groups in the acidic media used. It seems unlikely, however, that these phenomena could have happened to an extent of sufficient size to explain the results.

[In earlier experiments, maleic anhydride of a higher specific radioactivity was used, and the completeness of the demaleylation was checked by determination of the $^{14}$C-content of the protein peak after gelfiltration. Generally, almost complete removal was observed (> 90%). In this experiment a low specific activity was chosen, to prevent possible interference with the accurate determination of the incorporation of $^{14}$C-carboxymethylgroups. As a result, extent

of maleylation could be checked in this experiment, but not the efficacy of demaleylation, as this would have meant the sacrifice of an unacceptable proportion of the material. This strategy paid in so far that a good comparison was possible between the $^{14}C$-content of Tm-a and the CMCys-residues found during amino acid analysis (see V-II.3).]

V-II.1.2.1 Experimental details

Maleylation of atropinesterase. 2.5 µmol of atropinesterase was maleylated as described in III.1.3.1a. Specific radioactivity of 1,4-[$^{14}C$]-maleic anhydride was usually 40 µCi/mmol as in the experiment shown in fig.V.6; in a later study 10.8 µCi/mmol was used. The maleylated protein solution was desalted on a column (136 x 2.3cm) of Sephadex sf with 0.01 M Tris.HCl, pH 8.0, at a flow rate of 32 ml/h. Fractions of 5.3 ml were collected; their A280 and A250 were measured, and aliquots (1%) were used for LSC. Recovery of the maleylated protein was 92% according to the A280. The pool of maleylated protein was concentrated and denatured as decribed in III.1.3.1a.

Tryptic digestion of maleylated atropinesterase was performed as described in III.1.6.2. The digest was fractionated in two equal and identically treated portions on a Sephadex G-50 sf column (136 x 2.3 cm) eluted with 0.01 M Tris.HCl, pH 8.0, at 24 ml/h. Fractions of 2.4 ml were collected. Their A280 was measured and their $^{14}C$-content (LSC of 25 µl aliquots). Fractions 51-240 were analyzed in the PA (30 µl aliquots). Peak II was pooled as indicated in fig.V.6.

Demaleylation of pool II (see fig.V.6) and desalting of the demaleylated pool. Combined pools II were concentrated by lyophilization; acetic acid was added to obtain a solution of 200 nmol Tm-a per ml 33% acetic acid, which was incubated for 60 h at 40°C. Then acetic acid was added to a concentration of 50%. [$^{14}C$]-maleic acid was removed on a column (110 x 2.3 cm) of Sephadex G-50 sf in 50% acetic acid eluted at 27 ml/h. Fractions of 4.5 ml were collected; of nrs. 15-50 the A280 was measured, of fractions 45-104 the $^{14}C$ (LSC of 200 µl). The A280 peak was pooled (fractions 26-44) and lyophilized.

Reduction, carboxymethylation and desalting of Tm-a. 94.8% of the demaleylated peptide material was reduced and 2-[$^{14}C$]-carboxymethylated as described in III.1.3.2b. The modified peptide material was desalted as a solution in 50% acetic acid on the same column as used for the demaleylated peptide. Elution was with 50% acetic acid at 30 ml/h. Fractions of 3 ml were collected. Peptides were detected by measurement of the A280, by analyzing aliquots (0.4%) in the PA and by LSC for which 0.7% of each fraction was used. Peptide-containing fractions were pooled; leading and tailing edges of the peak were discarded. The remaining 5.2% of the demaleylated peptide material were carboxymethylated without reduction, as described in V-II.3.

Fractionation of 2-[$^{14}C$]-carboxymethylated Tm-a on DEAE-cellulose, pooling and removal of urea. 1.5 µmol Tm-a in 3.1 ml 0.01 M Tris.HCl, pH 8.5, containing 8 M urea, was chromatographed on a column (8.8 x

0.9 cm) of DEAE-cellulose (fig.V.7) equilibrated with buffer A, 0.01 M Tris.HCl, pH 8.0, containing 8 M urea. Elution was at roomtemperature, first with 4 ml A, then with A with a gradient from 0 to 0.05 M NaCl followed by a steeper gradient from 0.05 to 0.16 M, at a flow rate of 8.2 ml/h. The first gradient was made starting with 80 ml buffer A and at least 80 ml buffer B (B: A with 0.115 M NaCl). After 8 hours, indicated by the arrow in fig.V.7, buffer B was replaced by buffer A with 0.16 M NaCl, and elution was continued. Fractions of 1.64 ml were collected; their A280 and A225 (not shown) were measured, while aliquots (1.2%) of each fraction were removed for LSC. The gradients were made with the peristaltic pump as described in III.1.8.1b.
The peptide peaks a, b, c, and d, obtained after fractionation on DEAE-cellulose were freed of urea on the same Sephadex column as was used for the demaleylated peptide. Elution was performed with 50% acetic acid at 30 ml/h; fractions of about 5 ml were collected, while 1% of each fraction was used for LSC. The radioactive peak was pooled.

## V-II.2 DETERMINATION OF THE AMINO ACID SEQUENCE OF Tm-a

### V-II.2.1 DIRECT SEQUENCE DETERMINATION

When intact demaleylated and carboxymethylated Tm-a was subjected to the stepwise Edman degradation, the identity of the first 17 residues could be established as:

1 6 11 16

Gly-His-Gly-Gly-Thr-Ser-Ile-Pro-Lys-CMCys-CMCys-Tyr-Tyr-Val-Ser-Asp-Phe

By application of the manual dansyl-Edman procedure only the first 8 of these amino acids could be identified. This result was confirmed and extended to 17 residues in a sequencer analysis*. Fifteen of these residues were definitely characterized by HPLC-studies of the PTH-derivatives and of the regenerated amino acids, but the CMCys-residues at the positions 10 and 11 could not be identified unambiguously in this manner. Both PTH-derivatives gave two peaks in the HPLC-assay, in positions suggesting the presence of Asp + Glu (in the system used, PTH-CMCys co-chromatographs with PTH-Asp). The presence of CMCys at positions 10 and 11 in Tm-a followed from the relatively high $^{14}C$-content of the ethyl acetate extracts that contained the PTH-amino acids derived from the residues in these positions. (Further confir-

* Thanks are due to Dr. R. Amons and W.J.M. Pluijms of the University of Leiden for performance of the sequencer analyses.

mation of the sequence Cys-Cys at this place in the peptide will be given in subsequent sections).

The effectiveness of the sequencer analysis of the N-terminal sequence was reduced by the fact that the amount of peptide involved in the Edman degradation rapidly fell down, after a few steps, to about 10% of the initial quantity; this loss of material was probably caused by the poor solubility of the peptide in both acidic and basic aqueous solvents.

As elucidation of the complete primary structure of Tm-a by Edman degradation appeared impossible, selective fragmentation procedures had to be applied in order to obtain partial peptides suitable for further characterization of the amino acid sequence.

V-II.2.1.1 Experimental details

For the dansyl-Edman procedure 24 nmol Tm-a (pool a) was used. When His at position 2 was reached not only this residue but also part of the next residue was released. This phenomenon, which has already been reported by Blömback (13) and Schroeder (114), caused the appearance of the DNS-derivative of the next cycle beside the DNS-derivative of the position concerned. This effect of His at position 2, the normal background of preceding DNS-derivatives after several degradation steps -starting at position 8 in this experiment- and the limited amount of peptide available, made elucidation of only the first 8 residues possible with the manual method. The sequencer analysis was performed on 100 nmol Tm-a (pool b, purity > 90%). The 15 residues identified by HPLC all gave a clear peak at the position of the PTH-amino acid concerned. For Ile(7), Pro(8) and Lys(9) amounts of 14, 7 and 5 nmol were found. Identification of Ser(6), Ile(7) and Phe(17) was confirmed after conversion into the parent amino acid by HI. According to the $^{14}C$-radioactivity of the ethylacetate extracts that contained the PTH-amino acids derived from the positions 8, 9, 10, 11, 12 and 13, these extracts contained 0.3, 1.2, 6.7, 7.3, 3.5 and 1.7 nmol carboxymethyl groups, respectively.

V-II.2.2 SECONDARY CLEAVAGES WITH CNBr, TRYPSIN AND PEPSIN TO OBTAIN FRAGMENTS SUITABLE FOR SEQUENCE STUDIES

Tm-a contains 3 internal Met-residues and also 3 Lys. Cleavage with either CNBr or trypsin was expected, therefore, to yield 4 peptides, of a size suitable for sequence determinations if these residues were distributed over Tm-a not too irregularly. When both approaches were followed, also the overlapping sequences needed to place the partial

peptides in the correct order, would be obtained. In practice the CNBr-approach was not completely successful because the largest, N-terminal fragment (28 residues) could not be isolated. Also the attempts to isolate and characterize the tryptic peptides from Tm-a were hampered by technical problems. Prolonged incubation with trypsin was needed to attain well-advanced cleavage of the Lys-Asp and Lys-CMCys bonds present in Tm-a, which inevitably involved extensive occurrence of aspecific breaks. This phenomenon, together with the problem that some fragments gave rather diffuse or multiple peaks in various purification procedures, led to the result that either the final recovery of the products was poor or their purity was not very satisfying. Consequently the sequence of Tm-a could not be fully resolved on the basis of the data obtained with the CNBr-peptides and the tryptic peptides. Therefore, a suitable third method for the selective digestion of Tm-a was needed. Various enzymes were tried; pepsin appeared to give adequate results: a limited number of reasonably sized fragments in fair yields. Eventually, the peptic digest proved to be more useful for the elucidation of the total sequence of Tm-a than the results obtained with trypsin.

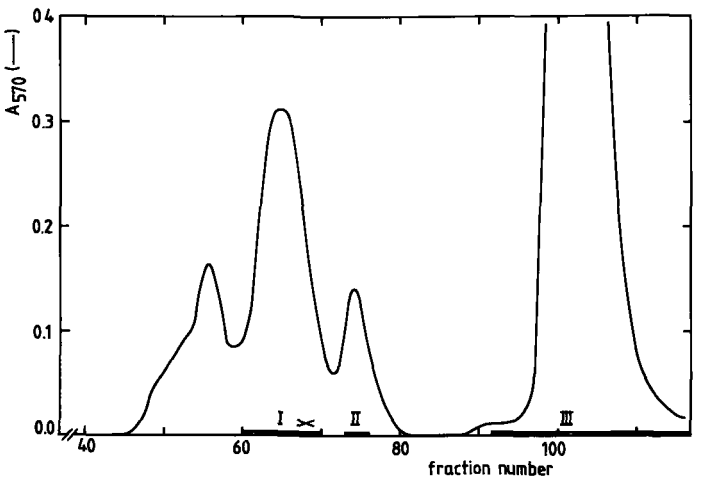



Fig.V.8. Fractionation of CNBr-peptides of Tm-a on Sephadex G-50 sf. The column was eluted with 33% acetic acid. Fractions of 1.4 ml were collected. Of the effluent 12% was analyzed for ninhydrin-positive material in the PA (A570:——). Fractions were pooled as indicated. High A570 in peak III is caused by excess CNBr eluted at this position. X: fractions 67, 68, 69.

V-II.2.2.1 CNBr-CLEAVAGE OF DEMALEYLATED Tm-a, ISOLATION OF CNBr-PEPTIDES AND DETERMINATION OF THEIR ACID SEQUENCE

Tm-a was digested with CNBr. Fractionation of the peptide mixture on Sephadex G-50 sf resulted in the profile shown in fig.V.8. The main peaks were pooled as indicated. Pool II contained a pure dodecapeptide, Tm-a CNBr-2 (nomenclature of peptides is defined in III.1.6). This fragment appeared to be identical to the active serine peptide CN-a isolated after CNBr-cleavage of atropinesterase, as judged by amino acid analysis and N-terminal sequence (table V.3). This result gave final proof of the presence of the active serine in Tm-a. Pool III contained the excess CNBr, which was removed by gelfiltration over Sephadex G-10, resulting in the isolation of pure Tm-a CNBr-3. Impure fractions of pool I (nrs. 67, 68, 69) were tentatively shown to contain the C-terminal fragment Tm-a CNBr-4 because of the high Arg-content. This was obtained in a homogeneous form by paperelectrophoresis of pool I. The N-terminal fragment Tm-a CNBr-1 could not be recovered from the eluate. Amino acid compositions, relative yields, N-terminal residues and the method used for final purification of CNBr-peptides of Tm-a are given in table V.3.

The sequence determinations were performed with the dansyl-Edman method. Investigation of CNBr-2 was not continued after it had been shown to be identical to CN-a. The sequence of the tripeptide CNBr-3 was fully resolved. In CNBr-4 (15 residues) the first 9 amino acids could be identified with the available amount (14 nmol). The results are included in the total survey of the structure of Tm-a in fig.V.12.

V-II.2.2.1.1 Experimental details

Cleavage of Tm-a (100 nmol) with CNBr: see III.1.7.7 and table III.1.

Fractionation of CNBr-peptides was performed on a column (55 x 2 cm) of Sephadex G-50 sf, eluted with 33% acetic acid at 16 ml/h. Of the effluent, 88% was collected in fractions of 1.4 ml, while 12% was used for on line detection of peptides in the PA (2.0 ml/h).

Fractionation of pool III was over a column (53 x 2 cm) of Sephadex G-10 eluted with 5% acetic acid at 14 ml/h; fractions of 1.6 ml were collected; 24% of the effluent was used for detection in the PA.

Fractionation of pool I was performed by HVE at pH 3.6 (90 min at 4000 V). Peptides were detected with fluorescamine, phenanthrenequinone and ninhydrin (see III.1.9.2) and eluted as described in III.1.8.2.

Table V.3 Amino acid composition of the CNBr-peptides of Tm-a.

| Amino acid | CNBr-2 | CNBr-3 | CNBr-4 |
|---|---|---|---|
| Asx | 1.37(1) | | 1.14(1) |
| Thr | 1.82(2) | | 0.96(1) |
| Ser | 1.07(1) | 0.98(1) | 1.80(2) |
| Glx | | | |
| Pro | | | 0.98(1) |
| Gly | 2.00(2) | 1.03(1) | 1.19(1) |
| Ala | 0.27 | | 2.21(2) |
| Val | 0.98(1) | | 1.85(2) |
| Met a) | 0.86(1) | 0.98(1) | |
| Ile | 0.85(1) | | 0.75(1) |
| Leu | 1.10(1) | | 1.18(1) |
| Tyr | | | |
| Phe | | | |
| His | 1.83(2) | | 0.87(1) |
| Lys | | | 1.21(1) |
| Arg | | | 0.86(1) |
| Trp | | | |
| No.of residues | 12 | 3 | 15 |
| Relative yield b) | 60% | 20% | 28% |
| $NH_2$-terminus | Gly-Leu-His | Gly | Thr |
| Purification c) | G-50 | G-50<br>G-10 | G-50<br>HVE |
| Pool(fig.V.8) | II | III | I |

Amino acid analyses were performed after 18 h hydrolysis; residues are reported as mol per mol of peptide, according to III.1.4.2. Figures in parentheses show the assumed integral values; for CNBr-3 these were equal to the values found by sequence analysis.

a) Met was determined as Hse.

b) The relative yield refers to the amount of peptide isolated after the last purification step compared to the amount of Tm-a degraded.

c) CNBr-peptides were fractionated by chromatography on Sephadex G-50 sf (G-50), possibly followed by Sephadex G-10 (G-10) or by paperelectrophoresis at pH 3.6 (HVE).

### V-II.2.2.2 DIGESTION OF DEMALEYLATED Tm-a WITH TRYPSIN, ISOLATION OF TRYPTIC (T-) PEPTIDES AND DETERMINATION OF THEIR AMINO ACID SEQUENCE

According to the presence of 3 Lys residues and the data known about the sequence of Tm-a mentioned in the preceding sections, ideally 4 tryptic peptides could be expected, one of 9 residues (T-1), one of 3-5 (T-4) and at least one (T-3) of fairly great length (> 26). However, when a tryptic digest was prepared and fractionated on Sephadex

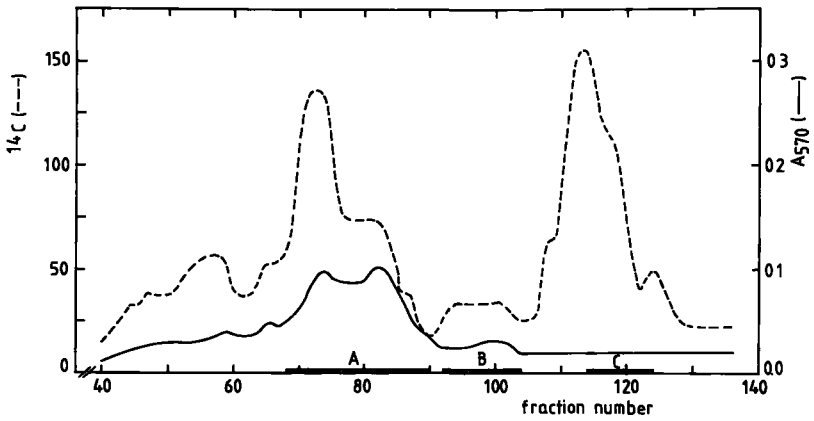



Fig.V.9. Fractionation of T-peptides of Tm-a on Sephadex G-50 sf. The column (107 x 2.3 cm) was eluted with 50% acetic acid. Peptides were detected on line with the PA (A570: ——). Fractions of 2.2 ml were collected. Aliquots were used for $^{14}C$-measurement by LSC, $^{14}C$ (---)is in cpm/0.1 ml. Fractions were pooled as indicated.

G-50 large peptides were almost absent (see fig.V.9). The fractions were pooled on the basis of amino acid analyses. Pool C contained a $^{14}C$-peptide consisting only of CMCys and Tyr, obviously representing the positions 10 through 13 of Tm-a (CMCys-CMCys-Tyr-Tyr), the N-terminal sequence expected for T-2; this tetrapeptide (T-2a) must result from an aspecific cleavage after Tyr. Pool B contained an impure Arg-containing fragment, probably Tm-a T-4. Fractionation of the first pool, A, by ion-exchange chromatography on Aminex A-5 (not shown) gave 4 homogeneous peptides, including the N-terminal fragment T-1. (the others were later identified as the peptides T-2b, T-3a and T-3c). The intact [$^{14}C$]-CMCys-containing fragment Tm-a T-2 was identified in impure fractions. The distribution of the radioactivity over many fractions, both after Sephadex and Aminex chromatography, and the kind of peptides isolated, indicated that -in addition to the normal tryptic cleavages- chymotryptic-like cleavages had occurred rather extensively. The yield of most peptides was insufficient for sequence elucidation.

In the subsequent experiments with trypsin, the digests were immediately fractionated on Aminex A-5. A representative elution profile of a tryptic digest is shown in fig.V.10. Four pure fragments of Tm-a, viz.

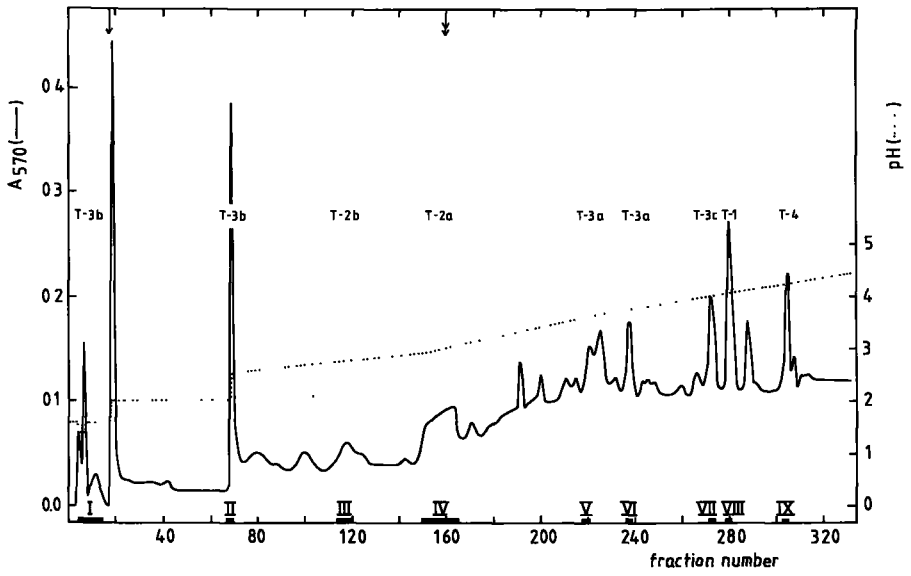



Fig.V.10. Fractionation of T-peptides of Tm-a by ion-exchange chromatography on Aminex A-5. The column (26 x 0.9 cm) was eluted at 50°C with pyridinium acetate buffers of a gradually increasing pH. The effluent was analyzed in the PA (A570: ——). The pH of the fractions (1.5 ml) was measured (pH: ...). ↓ ; ↓ : start first and second gradient, respectively. For details see text (V-II.2.2.2.1). Fractions were pooled as indicated.

T-1, T-2b, T-3c and T-4 were isolated. T-3a was recognized in two pools, V and IV; fractions of the latter contained almost homogeneous T-3a. A peptide composed only of neutral amino acids (later identified as T-3b) was found in the material that was not delayed by the column (pool I). This behaviour was quite suprising, as at this position in the eluate only strongly acidic peptides were expected. The peptide was purified by paperelectrophoresis (an additional amount of the same peptide was eluted at the pH-jump 2.0 to 2.5 (pool II)). The CMCys-containing tetrapeptide T-2a appeared present in pool IV but could not be obtained in a homogeneous form.

Together, the 7 peptides mentioned above represented the total composition of carboxymethylated Tm-a, but only 2 (T-1 and T-4) were the result of normal tryptic digestion; the others originated from aspecific (chymotryptic) breaks. Amino acid compositions, relative yields, N-terminal residues and the methods used for purification are given in table V.4.

Table V.4 Amino acid composition of the T-peptides of Tm-a.

| Amino acid | T-1 | T-2a[d] | T-2b | T-3a | T-3b | T-3c | T-4 | Total Tm-a |
|---|---|---|---|---|---|---|---|---|
| Asx | | 0.24 | 3.79(4) | 1.18(1) | 0.33 | 1.05(1) | | 6 |
| Thr | 1.01(1) | | 0.27 | 1.93(2) | 0.96(1) | | | 4 |
| Ser | 1.02(1) | 0.35 | 1.96(2) | 0.33 | 1.90(2) | 0.95(1) | 0.95(1) | 7 |
| Glx | | 0.26 | 1.20(1) | 0.24 | | | | 1 |
| Pro | 0.99(1) | | | | | 1.05(1) | | 2 |
| Gly | 2.87(3) | 0.47 | | 2.28(2) | 2.14(2) | 0.36 | | 7 |
| Ala | | 0.26 | 1.24(1) | 0.29 | 1.18(1) | 1.05(1) | | 3 |
| Val | | | 1.98(2) | 0.94(1) | 0.99(1) | | 0.95(1) | 5 |
| Met | | | | 0.92(1) | 1.48(2) | | | 3 |
| Ile | 1.01(1) | | 1.06(1) | 0.89(1) | | 0.93(1) | | 4 |
| Leu | | | 0.23 | 1.13(1) | 0.93(1) | | | 2 |
| Tyr | | 1.84(2) | | | | | | 2 |
| Phe | | | 1.82(2) | | | | | 2 |
| His | 1.03(1) | | | 1.93(2) | | 0.93(1) | | 4 |
| Lys | 1.07(1) | | 1.00(1) | 0.31 | | 1.05(1) | | 3 |
| Arg | | | | | | | 1.11(1) | 1 |
| Trp | | | | | | | | - |
| Cys a) | | 2.16(2) | | | | | | 2 |
| No.of residues | 9 | 4 | 14 | 11 | 10 | 7 | 3 | 58 |
| Relative yield b) | 17% | 21% | 15% | 13% | 18% (4%) | 24% | 41% | |
| $NH_2$-terminus | Gly | $^{14}$C-CMCys | Val | Met | Ser | Ala | Val | |
| Purification c) | Am. | G-50 | Am. | Am. | Am. + HVE | Am. | Am. | |
| Pool(fig.V.10) | VIII | C | III | V + VI | I + II | VII | X | |

Amino acid analyses were performed after 18 -24 h hydrolysis; see also legend to table V.2. Figures in parentheses show the assumed integral values or the values found by sequence studies.

a) Cys was determined as CMCys.

b) The relative yield of the peptides refer to the amount of peptide isolated after one purification step; the relative yield (of T-3b) in parentheses shows the amount of peptide after the final purification step, both yields are relative to the amount of peptide digested.

c) T-peptides were fractionated by chromatography on Aminex A-5 (Am.) or on Sephadex G-50 sf (G-50), possibly followed by paperelectrophoresis at pH 3.6 (HVE).

d) Peptides with contaminating amino acids or peptides.

Sequence studies. With some of these fragments ideal purity could not be reached, but in all cases where knowledge of the sequence was essential for the elucidation of the sequence of Tm-a, a satisfactory determination appeared possible. Not all tryptic peptides had to be sequenced completely as some sequences were already known from earlier experiments (sequencer results, CN-a and CNBr-cleavage of Tm-a). Furthermore, during these studies the outcome of the peptic degradation of Tm-a became available, which appeared superior to the results obtained with trypsin, so that a few more sequence determinations

could be omitted. The peptides concerned (T-1, T-3c and T-4) were characterized on the basis of their amino acid composition and the N-terminal sequences.

The sequence assumed for CMCys-containing T-2a was proven partly by manual Edman, partly by dansyl-Edman degradation.

Of the 14 residues of T-2b, 12 were identified by the dansyl-Edman procedure. The residue at position 9 was not determined, while after the penultimate residue (Asx) further sequencing failed. Attempts to liberate the C-terminal Lys with CPA + CPB were not successful either; no residues were released (table V.6), which was ascribed to the presence of Asp as the amino acid preceding Lys.

The undecapeptide T-3a consisted of the first 10 residues of CN-a, preceded by Met. It was characterized on the basis of the amino acid composition and the sequence of the first 5 residues.

Of T-3b, a decapeptide, 6 residues were sequenced. CPA released Val and Leu in equimolar amounts (table V.6).

The results obtained with the tryptic fragments of Tm-a are included in the survey in fig.V.12. Detailed information on CPA and CPB digestions is given in table V.6.

V-II.2.2.2.1 Experimental details

Tryptic digestion of demaleylated Tm-a. 130, 185 and 230 nmol Tm-a, respectively derived from a pool corresponding to pool b and from pools c and d of the DEAE-cellulose column (fig.V.7) were used for subsequent digestions with trypsin in 0.02 M Tris.HCl, pH 8.5, at a peptide concentration of about 250 nmol/ml (1.5 mg/ml); the initial solutions were not entirely clear. Digestion was started by the addition of trypsin; after 2 h of incubation at 37°C an equal amount of trypsin was added, to a final concentration of 0.22 mg/μmol peptide, and digestion was continued. After 16 h the tryptic digests still had not become clear. The reaction was stopped by adding acetic acid to a final concentration of 60%. Addition of a few drops of formic acid clarified the solution.

Fractionation of the first tryptic digest of Tm-a (130 nmol) was performed on a column (107 x 2.3 cm) of Sephadex G-50 sf eluted with 50% acetic acid at 19 ml/h. Fractions of 2.2 ml, were collected; 16% of the effluent (3.1 ml/h) was used for detection of peptides in the PA. Aliquots (4%) of each fraction were removed for LSC (fig.V.9). Pooled fractions were concentrated by lyophilization. Pools A and C contained 36% and 22%, respectively, of the eluted radioactivity; the relative yield of T-2a in pool C was 21%; attempts to isolate (more) pure peptide from the fractions with a high C-content by paperelectropho-

resis did not succeed. Pool B contained impure Tm-a T-4 with a relative yield of 22% on basis of the Arg-content.

Fractionation of tryptic digests of Tm-a on Aminex A-5. The material of pool A of the first digest, and the complete second and third digests were chromatographed on a column (26 x 0.9 cm) of Aminex A-5 at 50°C. Columns were developed with pyridinium acetate buffers given in table III.2. The tryptic digests were eluted at a flow rate of 16ml/h, first with 13 ml buffer A (0.025 M pyridine, pH 1.6), then 13 ml buffer B (0.025 M pyridine, pH 1.9); elution was continued by applying a two-steps gradient, the first from 200 ml buffer C (0.05 M pyridine, pH 2.4) to buffer D (0.2 M pyridine, pH 3.1), the second from 250 ml buffer D to buffer E (2 M pyridine, pH 5.0). When a pH of 4.6 was reached, the second gradient was replaced by 2 M pyridine. Gradients were made with a multichannel peristaltic pump. Fractions of 1.5 ml were collected; 20% of the effluent was used for detection of peptides in the PA. Of several fractions the pH value was measured. A representative elution profile is shown in fig.V.10. Fractionation of pool A was performed in essentially the same way; here the first gradient was made with 100 ml buffer C, the second with 180 ml buffer D, flow rate was 10.3 ml/h and aliquots of each fraction were used for LSC.

In general, the $^{14}C$-elution pattern was rather indistinct and little instructive how to pool; in later experiments, therefore, $^{14}C$-measurement was omitted.

Pool III, VII, VIII and IX (fig.V.10) contained the pure Tm-a fragments T-2b, T-3c, T-1 and T-4, respectively. The amount of impure T-2a in pool IV was too small for purification.

Position of T-3a. This peptide was eluted at 2 positions in the pH-gradient, at pH 3.57 and 3.72. At the lower pH (pool V), it was present in impure fractions, whereas pool VI contained almost homogeneous T-3a (relative yields 6% and 13%, respectively).

Purification of Tm-a T-3b from pool I was achieved by HVE at pH 3.6 for 120 min at 4000 V; peptides were detected by fluorescamine and eluted as described in chapter III. The amount of impure T-3b present in pool II was insufficient for further purification.

In almost all of the remaining impure fractions outside the pools mentioned, tryptic peptides of Tm-a could be tentatively identified, albeit in low yield. These peptides contained sequences which were already identified in pure fractions. In only one peak, pool IX, a peptide was recognized that could not be positioned in Tm-a. Evidently this fragment was derived from the main contaminating Tm-peptide present in pool II (fig.V.6) from which Tm-a was isolated.

Identification of T-1, T-3c and T-4. T-1 was recognized from its composition, presence of Lys and N-terminal Gly; T-3c because of its compositional correspondence with part of Tm-a CNBr-4, presence of Lys and N-terminal Ala, and T-4 from the presence of Arg.

Sequence determination of T-peptides was as described in chapter III.

For sequence determination of Tm-a T-2a ($CMCys_2$-$Tyr_2$) only 8 nmol was available (Sephadex G-50 sf, pool C; fig.V.9). An attempt to identify the first residue as DNS-derivative, with 1.5 nmol of peptide, failed. Evidently larger amounts of peptide were necessary for the identification of DNS-CMCys. Therefore, no attempt was made to identify the second residue as DNS-derivative. Instead, identification of the first

two residues was attempted by conversion of the thiazolinone derivative, present in the ethyl acetate extract, into the parent amino acid with 57% HI. This failed for the first CMCys because of technical problems, but succeeded for the second residue. Amino acid analysis indicated (per mol of Tm-a T-2a) 0.58 mol of "Ala", into which the thiazolinone of CMCys is converted, as was established in separate experiments. Confirmation of the presence of [$^{14}$C]-CMCys in the ethyl acetate extracts containing the thiazolinones from positions 1 and 2 was obtained by LSC; the $^{14}$C-determinations indicated the presence of 0.85 and 0.61 carboxymethyl groups, respectively, per mol of T-2a. Residues in positions 3 and 4 were identified as DNS-Tyr.

### V-II.2.2.3 DIGESTION OF DEMALEYLATED Tm-a WITH PEPSIN, ISOLATION OF PEPTIC (P-) PEPTIDES AND DETERMINATION OF THEIR AMINO ACID SEQUENCE

Pure Tm-a was digested with pepsin. The digest was fractionated by ion-exchange chromatography on Aminex A-5. The elution profile (fig.V.11) showed about 7 well-defined and separated peaks, in addition to one more diffuse peak-region. Fractions were pooled as indicated. According to amino acid analyses, pool III, V, VII and VIII

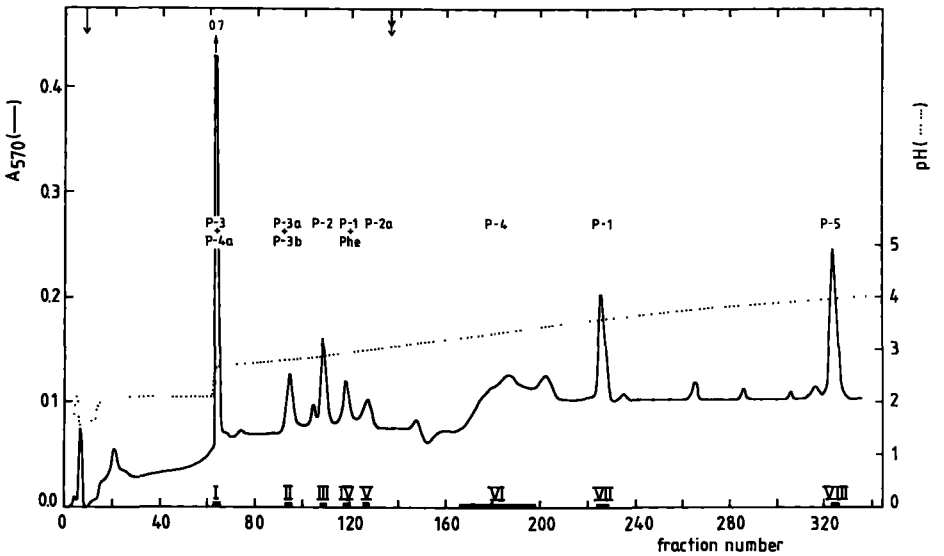



Fig.V.11. Fractionation of P-peptides of Tm-a by ion-exchange chromatography on Aminex A-5. The column (26 x 0.9 cm) was eluted at 50°C with pyridinium acetate buffers of a gradually increasing pH. Of the effluent 25% was fed into the peptide analyser (A570: ——). The pH of the fractions (1.40 ml) was measured (pH: ...). For details see text (V-II.2.2.3.1). Fractions were pooled as indicated.

contained (almost) pure peptides (table V.5), later identified as P-2, P-2a, P-1 and P-5, respectively. (P-2a is identical to P-2 but for one additional Phe at the C-terminus). P-1 was eluted also in peak IV, together with free Phe. Peak I comprised the heptapeptide P-3 and the pentapeptide P-4a, which could be separated and purified by paperelectrophoresis. Pool II showed exactly the same amino acid composition as P-3, but upon dansylation 2 N-terminal residues were found (Phe and Val). This pool appeared to consist of equimolar amounts of two fragments of P-3, resulting from partial cleavage after the fourth residue.

The residues of Tm-a not recovered with the above-mentioned peptides were found to be present in one large fragment P-4 (25 residues) that was eluted rather diffusely in the region of pool VI. Its presence was indicated by the amino acid analyses and the N-terminal Phe found in all fractions examined in this region. The spread-out elution is probably caused by adsorption of this large, rather hydrophobic peptide to the resin. After chromatography of pool VI on Sephadex G-50 sf, P-4 was recovered in a homogeneous form except for the presence of an almost identical "contaminating" fragment, i.e. P-4 foreshortened by one residue at the N-terminus. This "impurity", present in low yield, must be the result of partial cleavage of the N-terminal Phe of P-4, which also explains the loose Phe found in pool IV. Its identity was deduced from the low Phe-content of P-4 (table V.5) and the presence of a corresponding quantity of N-terminal Ile in the dansylation assay, next to Phe.

The amino acid compositions of all peptic peptides, covering all residues present in Tm-a, are given in table V.5 which also mentions relative yields, N-terminal residues and the methods for final purification.

P-1 (12 residues) was identified as the N-terminal fragment on the basis of amino acid composition and the sequence of 7 of the first 9 residues (dansyl-Edman procedure); residues in positions 3 and 6 were not identified to save material for further steps. Furthermore, the C-terminal residue was identified as Tyr, which was the only amino acid liberated by CPA (table V.6; this enzyme does not release CMCys,

Table V.5 Amino acid composition of the P-peptides of Tm-a.

| Amino acid | P-1 | P-2 | (P-2a)[d] | P-3 | P-4 | (P-4a) | P-5 | Total Tm-a |
|---|---|---|---|---|---|---|---|---|
| Asx | 0.37 | 1.10(1) | 1.11(1) | 1.97(2) | 2.17(2) | | 1.15(1) | 6 |
| Thr | 1.03(1) | | | | 2.89(3) | 0.65(1)[e] | | 4 |
| Ser | 1.02(1) | 0.94(1) | 1.02(1) | 0.91(1) | 1.80(2) | | 1.81(2) | 7 |
| Glx | 0.22 | | 0.26 | 1.06(1) | | | | 1 |
| Pro | 1.01(1) | | | | | | 1.07(1) | 2 |
| Gly | 2.86(3) | 0.20 | 0.46 | | 4.22(4) | 1.04(1) | | 7 |
| Ala | 0.24 | | 0.28 | 1.00(1) | 1.16(1) | 1.00(1) | 1.06(1) | 3 |
| Val | 0.22 | 1.00(1) | 0.91(1) | 0.96(1) | 1.89(2) | 0.97(1) | 1.01(1) | 5 |
| Met | | | | | 2.96(3) | | | 3 |
| Ile | 1.03(1) | | | | 1.85(2) | | 0.91(1) | 4 |
| Leu | | | | | 2.06(2) | 0.99(1) | | 2 |
| Tyr | 0.91(1) | 0.96(1) | 0.74(1) | | | | | 2 |
| Phe | | | 0.96(1) | 0.83(1)[e] | 0.76(1)[f] | | | 2 |
| His | 1.03(1) | | | | 2.08(2) | | 0.93(1) | 4 |
| Lys | 1.12(1) | | | | 0.97(1) | | 1.03(1) | 3 |
| Arg | | | | | | | 1.03(1) | 1 |
| Trp | | | | | | | | - |
| Cys a) | 1.62(2) | | | | | | | 2 |
| No.of residues | 12 | 4 | (5) | 7 | 25 | (5) | 10 | 58 |
| Relative yield b) | 22% | 31% | 11% | 61% (27%) | 31% (14%) | 25% (11%) | 28% | |
| $NH_2$-terminus | Gly | Tyr | Tyr | Phe | Phe(Ile) | Thr | Ala | |
| Purification c) | | | | + HVE | + G-50 | + HVE | | |
| Pool Aminex (fig.V.11) | IV+VII | III | V | I | VI | I | VIII | |

See legend to table V.2; for a), b) and d) see legend to table V.4. Peptides in parentheses contain duplicate sequences of other fragments; amino acid residues are to be ignored when the total of the fragments is computed.

c) Beside Aminex A-5, Sephadex G-50 (G-50) and paperelectrophoresis at pH 3.6 (HVE) were used for fractionation of P-peptides.

d) Peptides with contaminating amino acids or peptides.

e) Ratio number given for the N-terminal residues Phe in P-3 and Thr in P-4a is relatively low due to the reaction of the peptide with fluorescamine prior to hydrolysis.

f) The ratio number for Phe in P-4 is low due to the presence of a peptide identical to P-4 except for the absence of the N-terminal Phe.

the penultimate residue in P-1). The results fully agreed with the N-terminal sequence determined in intact Tm-a (V-II.2).

P-2, P-3, P-5 and P-4a (the C-terminal pentapeptide of P-4) were sequenced completely with the dansyl-Edman method, of P-2a only the N-terminal residue was determined. The nature of the Glx- and Asx-residues in P-2 and P-3 was established by degradation with CPA and/or LAP (see table V.6).

In P-4 the sequence of the first 7 residues was elucidated; the available amount did not allow more. Simultaneously with this determination, the sequence of the "contaminating" identical peptide, foreshortened by the N-terminal Phe, was recognized. Incubation of P-4 with CPA released equimolar quantities of Val and Leu (table V.6).
The sequence results obtained with the peptic fragments are summarized in fig.V.12.

V-II.2.2.3.1 Experimental details
Peptic digestion of Tm-a. 200 nmol Tm-a (pool a, fig.V.7) was digested with pepsin as described sub III.1.7.2 and in table III.1.
Fractionation of P-peptides was carried out as described for Tm-a T-peptides (V-II.2.2.2.1) on a column (26 x 0.9 cm) of Aminex A-5 at 50°C (fig.V.11). Elution was performed at a flow rate of 11 ml/h starting with 6 ml buffer A, followed by 6 ml buffer B and continued by a two-steps gradient going from 150 ml buffer C to buffer D and from 350 ml buffer D to buffer E. Fractions of 1.40 ml, were collected; 25% of the effluent was used for detection of peptides with the PA. Fractions were pooled as indicated.
Pure P-1 was present in pool VII with a relative yield of 22%; P-1 was eluted also in peak IV (relative yield 7%), together with Phe (relative yield 5%). The presence of free Phe was established as it was found both after C-terminal analysis with CPA (table V.6) and N-terminal analysis with dansylchloride, next to the C- and N-terminal residues of P-1, while amino acid analysis excluded the presence of a longer peptide beginning and ending with Phe.
Fractionation of pool I was carried out by HVE at pH 3.6 for 120 min at 4000 V, resulting in the isolation of P-3 and P-4a. Peptides were detected with fluorescamine (III.1.9.2.1a) and eluted as described in III.1.8.2. The ratio number given for Phe in P-3 and Thr in P-4a (table V.5) is relatively low because these N-terminal residues had reacted with fluorescamine used for the detection.
Purification of pool VI was performed on a column (107 x 2.3 cm) of Sephadex G-50 sf eluted with 50% acetic acid at 16 ml/h. Fractions of 2.2 ml, the remaining 20% of the effluent was used for detection in the PA. Ninhydrin-positive fractions, shown to contain P-4 by amino acid analyses and dansylation, were pooled.

## V-II.3 DETERMINATION OF THE PRESENCE OF AN S-S LINKAGE IN THE ACTIVE SERINE PEPTIDE Tm-a.

To establish whether the two carboxymethylcysteylgroups identified in Tm-a originated from two Cys residues connected via an S-S bridge to form one cystine, or were present as two cysteines with free SH-groups, the effect of omitting the reduction-step on the subsequent

Table V.6 Amino acids released by leucinaminopeptidase (LAP) or carboxypeptidase A and/or B (CPA+B) from CN-a and Tm-a and from T-peptides and P-peptides of Tm-a.

| Peptide | Enzyme | Incubation time (h) | Amino acids released[a] |
|---|---|---|---|
| | | | b) b) c) |
| CN-a | | | Gly-Leu-His-Asn-Thr-Thr-Val-Ile-Gly-His-Ser-Met |
| | LAP | 6 | 1.2 2.0 1.1 1.0 2.1 1.7 1.7 1.0 |
| Tm-a | | | ...-Lys-Val-Ser-Arg |
| | CPA/B | ½ | 1.0 1.0 1.0 |
| Tm-a T-2b | CPA/B | 18 | no amino acids found |
| Tm-a T-3b | | | Ser-Met-Gly-Ser-Met-Thr-Ala-Gly-Val-Leu |
| | CPA | 1 | 0.9 1.0 |
| | CPA | 18 | 0.4 0.4 0.5 0.3 1.0 1.0 |
| Tm-a P-1 | | | ...-Pro-Lys-CMCys-CMCys-Tyr |
| (pool VII) | CPA | 19 | 1.0 |
| (pool IV) | CPA | 19 | 0.8[d] |
| Tm-a P-2 | | | Tyr-Val-Ser-Asp |
| | CPA | 19 | 0.7 0.8 0.9 1.0 |
| Tm-a P-3 | | | Phe-Ala-Glu-Asp-Val-Ser-Asp |
| | CPA | 18 | 1.0 1.1 1.0 |
| | LAP | 18 | 0.8 0.9 0.9 2.0 1.0 1.4 |
| Tm-a P-4 | | | ...-Thr-Ala-Gly-Val-Leu |
| | CPA | 1½ | 1.0 1.0 |

Residues are ranked according to their sequence in the peptides.
a) Residues released are reported as mol of residue per mol of peptide.
b) Asn co-eluted with Ser.
c) Met was determined as Hse.
d) In pool IV 0.6 mol Phe per mol Tm-a P-1 was derived from free Phe.

reaction with iodoacetic acid was investigated. As was mentioned in V-II.1.2, the peptide mixture that contained Tm-a (pool II, fig V.6) was first demaleylated and then reacted with 2-[$^{14}$C]-iodoacetic acid, 94.8% following reduction with DTE, the remaining 5.2% without reduction. Subsequently, both portions were fractionated on Sephadex G-50. According to the $^{14}$C-content, all peptide-containing fractions of the unreduced and carboxymethylated material together contained 106 nmol [$^{14}$C]-carboxymethylgroups, whereas the corresponding fractions obtained with the reduced and iodoacetic acid treated material -after correction for the difference in starting quantities- had incorporated 2.3 x as much $^{14}$C (245 nmol). Evidently, reduction greatly increased the amount of carboxylmethylgroups bound to the peptide(s).

Of both eluates, fractions supposed to contain Tm-a (according to the

Kav- values; see V-II.1.2) were pooled in an identical manner and the amino acid compositions of these pools were compared. Although both pools appeared to contain one contaminating peptide in significant, be it equal quantities, the knowledge of the amino acid composition of Tm-a allowed the calculation of the amount of CMCys present per molecule of Tm-a, as well as the number of [$^{14}$C]-carboxymethylgroups. According to the amino acid analysis, 0.4 residues CMCys were present per molecule of non-reduced Tm-a, whereas after reduction 2.4 residues were found. $^{14}$C-determinations indicated the presence of 1.4 and 3.3 carboxymethylgroups, respectively. The increase in CMCys-residues as the consequence of reduction with DTE amounted to 2.0 residues per molecule of Tm-a, in excellent agreement with the increase of 1.9 carboxymethylgroups found with the $^{14}$C-determination. Evidently, reduction made two cysteine residues in Tm-a available for the reaction with iodoacetic acid. So, most probably the two cysteines identified in Tm-a were originally connected by an S-S bond to form one cystine. During sequence analysis of Tm-a the two cysteine residues were located next to each other. Model building of 2 neighbouring cysteine residues connected by an S-S bridge showed such a structure to be sterically possible, and without puting much constraint on the peptide chain.

The incorporation of 1.4 carboxymethylgroups per molecule of non-reduced Tm-a is not too suprising, as several groups are present in this peptide that can react with iodoacetic acid, which was used in excess ($\alpha$-and $\varepsilon$-$NH_2$ groups, the thioether bridge in Met and the imidazole in His). The 0.4 mole of CMCys found in the hydrolyzate of non-reduced Tm-a, however, was a bit unexpected. In reduced Tm-a, too, the amount of CMCys was 0.4 residue higher than the theoretical value. Obviously, reduction does not contribute to the formation of this apparent excess of CMcys. The exact nature of the product responsible for the high outcome of the CMCys analyses has not been identified. Tentatively it has been attributed to one of the products that are formed when S-carboxymethylmethionine is subjected to hydrolysis with HCl, viz. S-carboxymethylhomocysteine, which co-eluted with CMCys under the conditions used in our analyses.

V-II.3.1 Experimental details

Carboxymethylation and desalting of Tm-a. The modification of Tm-a, with or without reduction (already mentioned in V-II.1.2.1) with 2-[ C]-iodoacetic acid was as described in III.1.3.2. The 5.2% non-reduced modified peptide was desalted on a column (50 x 2.0 cm) of Sephadex G-50 sf, eluted with 50% acetic acid at 33 ml/h. Fractions of 3.1 ml were collected; 5% of the effluent was used for automatic detection of peptides with the PA. Of all fractions the A280 was measured and aliquots (6%) were used for LSC. Fractions were pooled on the basis of their Kav-values.

V-II.4 SEQUENCE ASSEMBLY

An unambiguous structure for the 58 residues-comprising tryptic peptide, Tm-a, that harbours the active serine of atropinesterase, can be constructed from the results presented sofar, which include the sequence of CN-a, the active peptide resulting from CNBr cleavage. A survey of the sequence assembly is shown in fig.V.12.

Sequencer analysis provides identification of the first 17 residues. The tryptic peptides T-1 and T-2a, and the peptic peptides P-1, P-2 and P-2a derived from this region are consistent with the sequencer data with respect to composition as well as to the residues positioned by sequence analysis.

T-2b, consisting of 14 residues, overlaps the N-terminal sequence of 17 by 4 residues, which extends the sequence of Tm-a to position 27. T-2b overlaps with P-2 and P-2a, includes P-3 and has a C-terminal overlap with the first 4 residues of P-4; this fixes the order in which these peptic fragments have to be ranked, and also gives final proof for the identity of the residues 9 and 14 of T-2b (positions 22 and 27 in Tm-a). The directly sequenced residues in P-4 extend the sequence of Tm-a to Leu in position 30; as for composition P-4 also covers the positions 31 through 48.

The first 3 residues of T-3a overlap the sequenced part of P-4, which determines the position of T-3a and the identity of the residues 31 and 32.

The structure of CNBr-2 is known because it is identical to the fully sequenced peptide CN-a. CNBr-2 overlaps two sequenced residues of P-4 and has its first 4 residues in common with the sequenced part of T-3a. (the amino acid composition of the remaining part of T-3a agrees

with the next 6 residues of CNBr-2). This fixes the position of CNBr-2 and extends the sequence of Tm-a to Met (40).

CNBr-2 and CNBr-3 must occupy internal positions in Tm-a; as CNBr-3 cannot be placed before CNBr-2, it should come between this fragment and CNBr-4. This identifies of the residues 41, 42 and 43.

According to its composition, T-3b should occupy a position at the C-terminus of P-4. This is in agreement with the fact that CPA liberates Leu and Val from both P-4 and T-3b. Further proof for the position of T-3b comes from the overlap with the last 2 residues of CNBr-2 and the subsequent 3 residues of CNBr-3 (which confirms the correct localization of CNBr-3). Both P-4 and T-3b overlap with the first 5 residues of CNBr-4. The sequence determined for these residues in CNBr-4 is consistent with all the data available from P-4 and T-3b concerning these positions: the 6th sequenced residue in T-3b coincides with the N-terminal residue of CNBr-4 and P-4a, the amino acid composition is in agreement and the last 2 residues, Val and Leu, are correct. Strong confirmation is provided by the composition and structure of the fully sequenced partial peptic fragment P-4a. Therefore, CNBr-4 is correctly linked with CNBr-3. This extends the primary structure of Tm-a to the last sequenced residues in CNBr-4, i.e. His (52).

The sequenced part of CNBr-4 overlaps with the first 4 residues of fully sequenced P-5, which provides the information for the remaining 6 positions in Tm-a. The sequence of P-5 is in complete agreement with the requirements put by the composition of CNBr-4, with the composition and the N-terminal residue of T-3c (and with its most probable C-terminal Lys), with the structure of T-4 and with the C-terminal sequence of intact Tm-a as revealed by treatment with CPA and CPB.

A few details remain to be discussed. The nature of the Asp at position 54 was established in an additional experiment, with newly isolated P-5. This batch was completely sequenced according to the manual Edman procedure, with results confirming the earlier data. Asp (54) was identified as the PTH-residue. The presence of Asp at position 26

Fig.V.12 Amino acid sequence of Tm-a and CN-a, fragments containing the active serine residue.

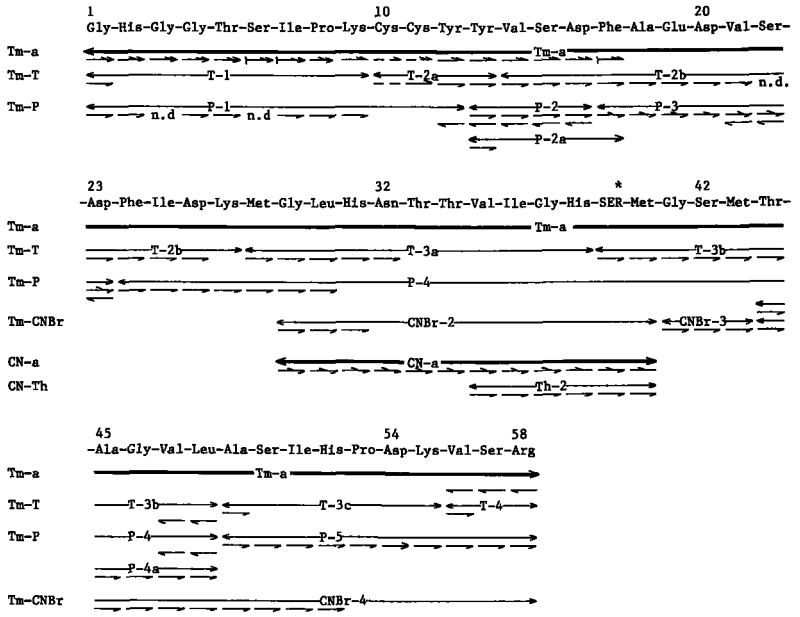



Fig.V.12. Tm-a was derived by tryptic digestion of maleylated atropinesterase, CN-a by CNBr-cleavage of the unmodified protein. Peptides formed by secondary cleavage of Tm-a are indicated as Tm-T, Tm-P and Tm-CNBr for products obtained with trypsin, pepsin and CNBr, respectively. Similarly, CN-Th indicates thermolysin digested CN-a.

- - Residues which were identified by $^{14}C$-determination of CMCys.
—→ Determined by dansyl-Edman degradation.
⊢— Identified by amino acid analysis after back hydrolysis of PTH-derivatives with HI.
—≫ Determined as PTH amino acid by HPLC after sequencer analysis.
—→ Determined as PTH amino acid by thin layer chromatography.
—→ Residues released by LAP.
←— Residues released by CPA or CPB or both.

rather than Asn was suggested on the basis of the failure of CPA+CPB to remove the C-terminal Lys from T-2b. In other experiments in which atropinesterase was fragmented, the peptide Ile-Asp-Lys-Met was isolated; in this fragment the Asp was positively identified as PTH-residue.

## V-II.5 DISCUSSION

A fairly long fragment of atropinesterase was sequenced, which comprised the active serine at a position sufficiently remote from either end (38 and 19 residues) to allow a sensible comparison with other serine-hydrolases. Very little, if any, sequence homology could be observed, however, with the corresponding sequences in hydrolases related to prothrombin or in the subtilisins. Occasionally a few residues coincided (e.g. -Gly- X -SER*- X -Gly-Ser- with the α-lytic protease, or -Gly- X -SER*-Met- X -Ser- with the subtilisins) or certain sequence similarities were noted (e.g. -His-Pro-Asx- 13 residues beyond SER* in atropinesterase, and 17 in subtilisin), but when an alignment was attempted no significant result could be obtained.

The absence of any distinct sequence homology did not exclude the existence of a distant evolutionary relationship between atropinesterase and either of the two superfamilies of serine-hydrolases. It appeared advisable, however, to base attempts to demonstrate such a remote sequence homology on a comparison of total sequences rather than of active serine peptides. It was decided, therefore, to continue the sequencing of atropinesterase in order to elucidate the complete primary structure. A description of the subsequent studies on a possible relationship with other proteins, including a computer search for sequences related to that of the 24 residues surrounding the active serine, is given in chapter VIII.

## CHAPTER VI ISOLATION AND CHARACTERIZATION OF THE TRYPTIC PEPTIDES OF MALEYLATED ATROPINESTERASE AND OF FRAGMENTS FORMED BY CLEAVAGE WITH CNBr OR PEPSIN

Atropinesterase comprises 7 Arg residues; digestion with trypsin after maleylation, therefore, was expected to yield 8 peptides, with an average length of 34 residues, 7 containing a C-terminal Arg and 1 a Lys, the C-terminal residue of the protein. Several tryptic degradations were carried out. In the initial experiments, extensive aspecific cleavage occurred, of an apparent chymotryptic character although TPCK-treated trypsin had been used. By changing conditions and by using different preparations of trypsin, this problem could be largely overcome though not completely. One of the attempted changes involved the use of clostripain (63), a protease supposed to cleave exclusively peptide bonds at the C-terminal side of Arg, but this enzyme did not give better results than trypsin. Under the conditions developed, a reproducible and fairly specific digestion was obtained, resulting in the formation of all 8 Tm-peptides in good yields. The fractionation of these digests was worked out in numerous experiments, which led to the more or less general procedures described in VI-I. In the description of the isolation of the Tm-peptides in a pure state, the results of a few experiments have been combined, as all peptides were isolated more than once, with sometimes slight modifications in the methods used. The Tm-peptides are numbered according to the sequence in which they occur in the intact protein. The purification of Tm-5 (=Tm-a), the active serine-containing fragment is not described here, as this isolation was presented in chapter V.

After CNBr-cleavage of atropinesterase, which was expected to give 10 peptides, only 6 CN-fragments (including the active serine peptide CN-3) comprising 82 residues, could be isolated in pure state (VI-II). Furthermore, one tripeptide was identified in impure fractions in low yield. The remaining fragments were present in a mixture of large CN-peptides ("CN-"); pure peptides could not be isolated from this mixture despite the many fractionation methods attempted. When it was

attempted to deduce the primary structure of atropinesterase on the basis of the sequences of the Tm- and CN-peptides (chapter VII) it became clear that additional information was needed. To obtain the necessary data a number of peptic fragments of the protein were studied. The isolation and the characterization of these P-peptides are described in VI-III.

## VI-I ISOLATION, PURIFICATION AND CHARACTERIZATION OF ALL TRYPTIC (Tm-) PEPTIDES OTHER THAN THE ACTIVE SERINE CONTAINING Tm-5

### VI-I.1 FRACTIONATION OF THE TRYPTIC AND CLOSTRIPAIN DIGESTS OF THE MALEYLATED PROTEIN

a. In a typical experiment 2.5 µmol of 1,4-[$^{14}$C]-maleylated atropinesterase was digested with trypsin, resulting in hydrolysis of all Arg-bonds and a few aspecific breaks, according to tests with CPA and CPB.

Half of the tryptic digest was chromatographed on Sephadex G-50 sf. The elution profile is shown in fig.VI.1, which is identical to fig.V.6. Fractions were pooled as indicated. Amino acid analysis and measurement of radioactivity showed the distribution of peptide material and of $^{14}$C-radioactivity over the different pools indicated in fig.VI.1. The amount of [$^{14}$C]-maleylgroups corresponded fairly well with the Lys content of each pool; only pool V contained much more $^{14}$C than expected, presumably due to maleic acid not completely removed when the maleylated protein was purified by gelfiltration. From pools I, III, IV and V, 7 Tm-peptides were isolated after application of the purification procedures given in scheme VI. The 8th fragment, the already isolated and characterized active serine peptide Tm-5, was present in pool II (see chapter V). Pools a, b and c, indicated by a broken bar, did contain small amounts of amino acids; no distinct peptides were recognized.

b. Clostripain was used for digestion of atropinesterase in one experiment. Despite the claimed specificity of the enzyme for

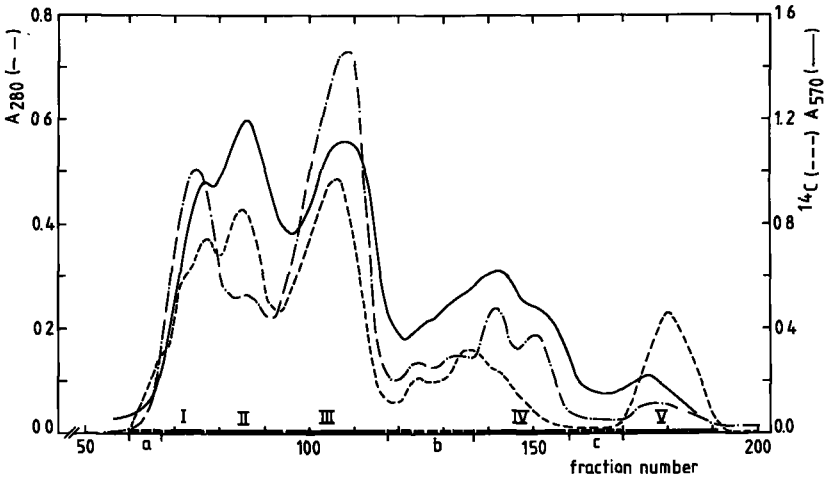



Fig.VI.1. Fractionation of Tm-peptides from 1.2 µmol 1,4-[$^{14}$C]-maleylated atropinesterase on Sephadex G-50 sf. The column (136 x 2.3 cm) was eluted with 0.01 M Tris.HCl, pH 8.0, at 24 ml/h. Fractions of 2.4 ml were monitored at 280 nm (-·-·-). Aliquots (1.3%) were analyzed in the PA (A570:——); aliquots (1.0%) were removed for LSC, $^{14}$C (---) is in $10^3$ cpm/50 µl. Fractions were pooled as indicated by solid bars (I-V) and broken bars (a-c).

Percentages of $^{14}$C-radioactivity and the amounts of amino acids present in each pool are given in percentages relative to the amounts applied to the column.

| Pool | a | I | II | III | b | IV | c | V |
|---|---|---|---|---|---|---|---|---|
| % amino acids | 1.3% | 12.4% | 23.0% | 22.2% | 6.3% | 10.0% | 1.2% | 2.8% |
| % $^{14}$C-radioactivity | 2.7% | 11.4% | 21.9% | 29.9% | 6.9% | 7.3% | 0.4% | 11.6% |

Arg-bonds, maleylation appeared necessary, as without this protection breaks occurred at the C-terminus of Lys-residues. In the maleylated protein all Arg-bonds were broken and very little aspecific breaks were induced. The elution profile on Sephadex G-50 sf (not shown) was essentially the same as that of the Tm-peptides of maleylated atropinesterase (fig.VI.1) and fractions were pooled similarly.

VI-I.1.1 Experimental details

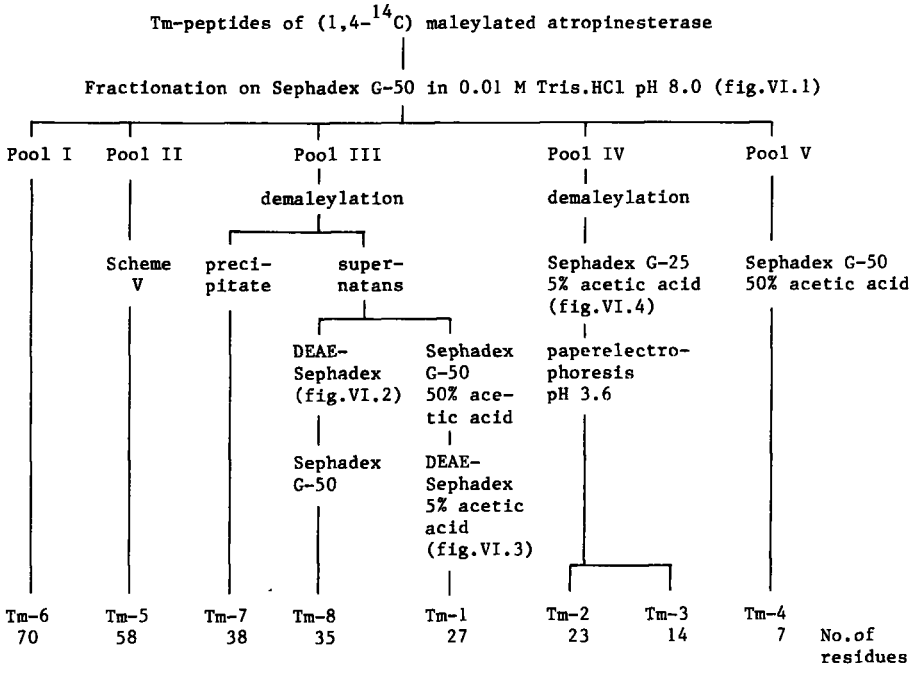
Maleylation of atropinesterase and tryptic digestion of the maleylated protein were carried out as described in III.1.3.1a and III.1.6.2. Fractionation of tryptic peptides was performed on a column (136 x 2.3 cm) of Sephadex G-50 sf eluted with 0.01 M Tris.HCl, pH 8.0, at 24 ml/h (see V-II.1.2.1). Fractions were pooled as indicated. After chromatography, 92% of the total $^{14}$C-radioactivity and 79% of the total amount of amino acids applied to the column were recovered in the 8 pools, comprising fractions 55-195 (amino acid content of the pools

was measured by amino acid analysis after hydrolysis with 6 N HCl for ca.24 h). Results are shown in fig.VI.1.

Digestion of atropinesterase with clostripain. Digestion of non-maleylated atropinesterase with clostripain caused extensive cleavage at the C-terminal side of Lys-residues, as incubation of the digest with CPB released large amounts of Lys, next to Arg. Gelfiltration on Sephadex G-50 sf showed many small fragments to be present.
Subsequently 0.84 µmol (25 mg) maleylated atropinesterase was digested, at a protein concentration of 115 nmol/ml in 0.08 M Tris.HCl, pH 7.7, with 3.6% (w/w) clostripain (0.9 mg/ml) for 5 h at 25°C. Clostripain (80 Units) was preincubated in 2.5 ml 0.05 M potassium phosphate, 1mM dithiothreitol, pH 7.6, with 0.28 ml soybean trypsin inhibitor (1 mg/ml phosphate buffer, pH 7.6) for 2 h at room temperature. Specificity of the digestion was checked with CPB and CPA. With CPB all Arg-residues were released, with CPA only small quantities of amino acids were liberated.

Fractionation of Cm-peptides was performed on a column (128 x 2.3 cm) of Sephadex G-50 sf eluted with 0.01 M Tris.HCl, pH 8.0, at 7.6 ml/h; fractions of 2.8 ml were collected; aliquots (1.5%) were used for detection of peptides in the PA. Aliquots (1.3%) were used for LSC.

Scheme VI Isolation of tryptic(Tm-) peptides of maleylated atropinesterase.



## VI-I.2 ISOLATION OF THE Tm-PEPTIDES FROM THE DIFFERENT POOLS

### VI-I.2.1 ISOLATION AND CHARACTERIZATION OF Tm-6 AND Cm-6a FROM POOL I

a (Tm-6): Amino acid analyses of aliquots of the fractions of peak I (fig.VI.1) showed a constant composition over the first half of the peak. These fractions were combined in pool I. From the amino acid composition (table VI.1) the minimum number of residues was calculated on the assumption that only one peptide was present (in these calculations, Arg and possibly other amino acids present in relative low amount were assumed to occur only once in the peptide). The result agreed rather well with the approximate molecular weight roughly estimated from the position of the fractions in the eluate (Kav-values). This suggested the presence of one fairly pure peptide (Tm-6) of about 70 residues.

Homogeneity was confirmed by endgroup determination. Reaction with dansylchloride showed Leu to be the only N-terminal residue and 3 subsequent steps of dansyl-Edman degradation revealed only one amino acid per step, showing the N-terminal sequence Leu-Val-Leu-Ile. After incubation with CPA and CPB, Arg and a few other C-terminal residues were identified in stoichiometric quantities (table VI.2); for unknown reason only 0.5 equivalent Ser was found; later experiments with the demaleylated peptide resulted in the liberation of Lys and the same 8 residues with better equivalency for Ser as also shown in table VI.2.

[In most experiments some heterogeneity was observed in the peptide material of pool I at the C-terminal side. In all cases, Leu was shown to be the only N-terminal residue, but amino acid analyses repeatedly showed low stoichiometric values for Arg compared to the integral numbers assumed for pure Tm-6; this phenomenon was -usually- also observed for some residues (Ser and Ala) preceding Arg. Furthermore, in these preparations CPA and CPB released less than one mol of Arg and of some other residues per mol of peptide, whereas other residues were liberated in equivalency to the amount of peptide. These observations suggested that some aspecific cleavage had occurred of one or more peptide bonds near to the C-terminal

side of Tm-6. This notion was supported by the fact that in these preparations small amounts of amino acids were released by CPA alone, indicating that a fraction of the peptide did not contain a C-terminal Arg. The amount of each residue released appeared to depend on the extent to which aspecific cleavage close to the C-terminal side had occurred.]

b (Cm-6a): After fractionation of the clostripain (Cm-) digest on Sephadex G-50 sf a pool I was made, strictly analogous to pool I of the Tm-peptides (fig.VI.1) by combining fractions with the corresponding Kav-values. Amino acid analysis (not shown) gave results very similar to those of Tm-6. Most of the residues of Tm-6 were present, and the N-terminal sequence Leu-Val-Leu-Ile, as determined by dansyl-Edman degradation, was identical. However, the ratio numbers of Trp, Ala, Val, Leu, His and Arg, the C-terminal residues of Tm-6 released with CPA and CPB, were low compared to those found in hydrolysates of Tm-6. Evidently, this pool contained a fragment of Tm-6 shortened at the C-terminal side by some residues (Cm-6a) together with a small amount of Cm-6, a peptide identical to Tm-6. This result strongly resembled the observations made in some degradation experiments with trypsin (preceding section). (Later experiments involving tryptic digestion of demaleylated material confirmed the identity of Cm-a: all tryptic peptides of Tm-6 were formed except the C-terminal fragment; instead a shorter peptide identical to the N-terminal part of the fragment was found.)

VI-I.2.1.1 Experimental details

Incubation with CPA and CPB of Tm-6 and of all other Tm-peptides discussed in this chapter, was performed as described in chapter III. Details are given in table VI.2.

During dansyl-Edman degradation of Tm-6 beside Val on position 2 also DNS-Val-Leu was detected; its position on a polyamide sheet is given in III.1.10.1b.

Determination of the amino acid composition of Tm-6. Hydrolysis of samples was performed for 24, 48 and 72 h at 110°C. Hydrolysis for 48 and 72 h slightly increased the amount of Leu compared to the value found after 24 h, the ratio number going from 8.24 to 8.56. (This increase was not due to more extensive bond splitting of Leu/Val or Leu/Ile bonds as constant values were found for Ile and Val; possibly

other peptide bonds were involved, for instance Leu-Lys, as was suggested by the small increase also detected for Lys after longer hydrolysis times).

Desalting of pools I, containing Tm-6 and Cm-6a, respectively, was performed on a column (62 x2 cm) of Sephadex G-15 eluted with 1% collidinium acetate, pH 8.3, at 21 ml/h. Fractions of 3.5 ml were collected; peptide material was detected by LSC of aliquots (0.7%) of each fraction and by measurement of the A280. Peptide containing fractions were pooled and concentrated by lyophilization.

Demaleylation of Tm-6 and Cm-6a was performed as described in III.1.3.1b at a concentration of ca. 150 nmol/ml. The peptide precipitated during acidification but the solution became clear after incubation for 66 h at 40°C. The demaleylated peptide was concentrated or dried by lyophilization.

Table VI.1 Amino acid composition of tryptic (Tm-) peptides of maleylated atropinesterase including Tm-5, described in chapter V.

| Amino acid | Tm-1 | Tm-2 | Tm-3 | Tm-4 | Tm-5 | Tm-6 | Tm-7 | Tm-8 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Asx | 4.03(4) | 3.18(3) | 1.07(1) | 1.14(1) | 6.07(6) | 7.22(7) | 4.71(5) | 4.12(4) | 31 |
| Thr a) | | 2.04(2) | | | 3.81(4) | 4.22(4) | 4.10(4) | | 14 |
| Ser a) | 1.96(2) | 1.96(2) | 2.72(3) | | 6.70(7) | 3.03(3) | 3.21(3) | 1.11(1) | 21 |
| Glx | 3.11(3) | 1.16(1) | | | 1.42(1) | 6.25(6) | 4.00(4) | 3.93(4) | 19 |
| Pro | 1.95(2) | 1.94(2) | 1.01(1) | | 2.21(2) | 4.95(5) | 1.24(1) | 2.23(2) | 15 |
| Gly | | 3.10(3) | | | 6.68(7) | 3.42(3) | 1.86(2) | 2.01(2) | 17 |
| Ala | 4.13(4) | 0.26 | 2.11(2) | 1.19(1) | 3.25(3) | 7.96(8) | 3.35(3) | 4.05(4) | 25 |
| Val | 1.10(1) | 1.93(2) | | | 5.21(5)+ | 6.83(7) | 0.36 | 1.16(1) | 16 |
| Met b) | 1.00(1) | | | | 3.09(3) | 1.08(1) | 2.11(2) | 2.00(2) | 9 |
| Ile | 4.61(5)* | | | | 3.78(4)+ | 1.30(1) | 3.71(4)x | 1.23(1) | 15 |
| Leu | | 3.05(3) | 1.91(2) | 2.82(3) | 2.26(2) | 8.56(9)+x | 3.25(3) | 2.97(3) | 25 |
| Tyr | | 1.75(2) | | 0.90(1) | 1.95(2) | 1.07(1) | 1.13(1) | 0.89(1) | 8 |
| Phe | | | 1.87(2) | | 1.98(2) | 2.16(2) | 1.31(1) | 3.57(4) | 11 |
| His | | 1.10(1) | | | 3.79(4) | 2.10(2) | | 0.98(1) | 8 |
| Lys | 2.86(3) | 0.30 | 1.98(2) | | 3.14(3) | 6.82(7) | 2.17(2) | 3.81(4) | 21 |
| Arg | 0.99(1) | 0.98(1) | 0.95(1) | 0.97(1) | 0.94(1) | 0.80(1) | 0.91(1) | | 7 |
| Trp b) | 0.97(1) | 0.84(1) | | | | 3.18(3) | 2.10(2) | 0.93(1) | 8 |
| Cys c) | | | | | 1.79(2) | | | | 2 |
| No. of residues | 27 | 23 | 14 | 7 | 58 | 70 | 38 | 35 | 272 |
| Yield | 27% | 18% | 18% | 28% | 43% | 24% | 46% | 24% | |
| $NH_2$-terminus | Glx | Tyr | Ala | Tyr | Gly | Leu | Gly | Ala | |

The procedures used for the isolation of these peptides are summarised in schema VI. Residues were determined after 24 h hydrolysis of ca. 10 nmol of peptide and in some cases also after 48 (+), 72 (x) or 96 h (*) h hydrolysis with 6N HCl at 110°C and are reported as mol per mol peptide (see III.1.4.2). Values are the mean of at least 3 amino acid analyses. Figures in parentheses show the nearest integral number (being equal to the number of each residue found by sequence studies (Chapter VII)).
Yields refer to degraded amount of atropinesterase.
a) Values for Thr and Ser after extrapolation to zero time, except for Tm-2 and Tm-8.
b) Met and Trp were determined after hydrolysis in the presence of 2% TGA
c) Cys was determined as CMCys.

(+), (x) and (*) Value for Val, Ile or Leu taken from the sample hydrolysed for 48, 72 or 96 h, respectively because this value was significantly higher than from the 24 h hydrolysate.

VI-I.2.2 ISOLATION AND CHARACTERIZATION OF Tm-1, Tm-8, AND Tm-7 FROM POOL III

Pool III (fig.VI.1) appeared to comprise a number of peptides. After demaleylation an acid-soluble and an acid-insoluble fraction were obtained. The acid-soluble fraction was chromatographed on DEAE-Sephadex, resulting in two well-separated peaks (fig.VI.2). According to amino acid composition, endgroup determination and the roughly estimated molecular weight, the first peak contained a practically pure peptide, comprising 27 residues, Tm-1, in a yield of 51%. According to the same criteria the second peak, pool 2, contained a somewhat less pure peptide of 35 residues, Tm-8, in a yield of 48%. Both peptides were sufficiently pure for sequence studies (chapter VII), but in some experiments additional purification steps were performed.

Best results with Tm-1 were obtained when the acid-soluble fraction, resulting after demaleylation of pool III (fig. VI.1), was first chromatographed on Sephadex G-50 sf. The second of 2 partially separated peaks was then subjected to chromatography on DEAE-Sephadex.

For the purification of Tm-8, pool 2 of the DEAE-Sephadex eluate (fig.VI.2) was fractionated on Sephadex G-50 sf which yielded 3 peaks (fig.VI.3). Pool I contained Tm-8 in pure state with the amino acid analysis shown in table VI.1, in an over-all yield of 24%. The absence of Arg suggested Tm-8 to be the C-terminal peptide of the protein.

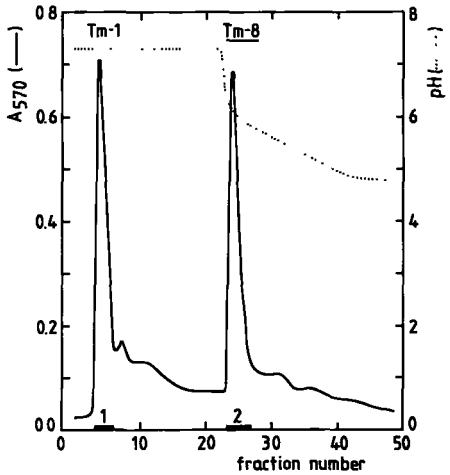



Fig.VI.2. Separation of Tm-1 and Tm-8 on DEAE-Sephadex A-25. The acid soluble fraction of the demaleylated pool III (fig.VI.1) was chromatographed on a column (20 x 0.9 cm) of DEAE-Sephadex A-25, eluted at 30 ml/h with an acetate buffer, pH 7.3, followed by a pH-gradient and finished by washing with 95% acetic acid. Of the effluent 10% was used for on line detection of peptides in the PA (A570:——). The pH of the fractions (2.3 ml) was measured (pH:...). Fractions were pooled as indicated.

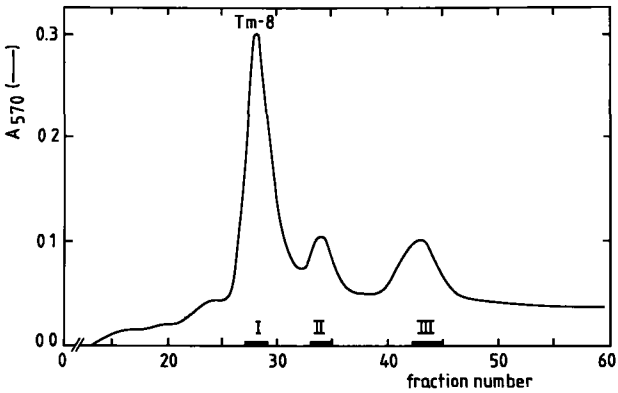



Fig.VI.3. Purification of Tm-8 on Sephadex G-50 sf. Pool 2 (fig.VI.2) was fractionated on a column (150 x 0.9 cm) of Sephadex sf eluted with 5% acetic acid at 13.3 ml/h. Fractions of 1.3 ml were collected; 25% of the effluent was used for on line detection of peptides in the PA (A570:——). Fractions were pooled as indicated. Pool I contained pure Tm-8.

Digestion with CPA and CPB confirmed this notion because the amino acids released by these enzymes (table VI.2) were identical to the residues liberated from the C-terminus of atropinesterase (table IV.3). Pool II (fig.VI.3) appeared to contain 2 fragments of Tm-8, Tm-8a and Tm-8b, respectively, as the amino acid composition was practically identical to that of Tm-8, whereas sequence determination of the first 3 residues indicated the presence of 2 N-termini, one with the same sequence Ala-Ala-Leu as found for Tm-8 and the other with the sequence Pro-Glu-Met. For either peptide a yield of about 5% was calculated. In pool III no peptides were recognized.

In the acid-insoluble material obtained after demaleylation of pool III (fig. VI.1) a fairly pure peptide, comprising 38 residues (Tm-7), appeared to be present, according to amino acid composition (table VI.1), the roughly estimated molecular weight and endgroup determination; yield was 46%. Several steps of dansyl-Edman degradation confirmed homogeneity. Incubation with CPA and CPB for different times resulted in the liberation of 5 or more residues from the C-terminus, as shown in table VI.2.

Table VI.2 Amino acids released by carboxypeptidase A and B (CPA/B) or leucin-aminopeptidase (LAP) from Tm-peptides of maleylated atropinesterase

| Peptide | Enzyme | Incubation time | Amino acids released [a] |
|---|---|---|---|
| Tm-2 | | | ...-Pro-Val-Leu-Leu-Leu-His-Gly-Tyr-Thr-Asp-Thr-Ser-Arg |
| | CPA/B(B) | 18 h | 0.4 1.7 0.6 0.8 0.6 0.6 1.3 1.0 1.0 |
| Tm-3 | | | Ala-Phe-Ser-Ser-Leu-Ala-Pro-Phe-Leu-Ser-Lys[c]-Asp-Lys-Arg |
| | LAP | 15 min | 2.3 2.0 2.4 2.0 1.0 n.d. 0.8 1.0 |
| | | 4 h | 0.9 1.0 1.5 1.1 |
| Tm-4 | | | Tyr-Leu-Ala-Leu-Asp-Leu-Arg |
| | CPA/B(B) | 18 h | 0.8 1.2 1.0 3.0 1.0 |
| | " (S) | 16 h | 0.3 0.5 0.4 1.7 1.0 |
| Tm-5 | | | ...-Asp-Lys-Val-Ser-Arg |
| | CPA/B(S) | 30 min | 1.0 1.0 1.0 |
| Tm-6 | | | ...-Pro-Lys-His-Val-Trp-Leu-Ser[b]-$(Ala)_2$-Arg |
| | CPA/B(S) | 30 min | 0.9 1.1 1.0 1.1 0.5 2.0 0.9 |
| (dem.) | CPA/B(B) | 20 h | 0.6 0.8 0.9 0.7 0.8 0.7 2.0 1.0 |
| Tm-7 | | | ...-Pro-Met-Thr-Glu-Ser-Met-(Gln+Asn)[d]-Asp-Ile-Arg |
| (dem.) | CPA/B | 5 min | 0.4 0.2 0.4 0.4 |
| | | 12 min | 0.3 0.3 0.7 1.6 0.7 1.0 1.0 |
| | | 18 h | ? 0.6 0.8 1.1 1.7 0.8 0.9 1.0 |
| Tm-8 | | | ...-Pro-Glu-Met-Val-Ala-Lys-Asp-Leu-Asn[e]-Glu-Phe-Leu-Lys |
| (dem.) | CPA/B(B) | 18 h | 1.0 1.0 1.2 1.1 0.7 1.0 1.1 1.9 2.0 |

Peptides were incubated at 37°C for the period indicated according to the conditions described in III.1.5. To facilitate interpretation, the results are presented in combination with the later determined C-terminal sequences of these peptides.

a) residues released, as determined by automated amino acid analysis, are given as mol per mol peptide.
b) No explanation can be given for the low value of Ser sometimes found.
c) Lys co-eluted with $NH_3$, making calculation impossible.
d) Asn and Gln eluted at the position of Thr, but were recognized from the high A440/A570 ratio; they may have masked some Thr released in the 18 h sample.
e) Asn eluted at the position of Ser, but was recognized from the high A440/A570 ratio.

(B)=Boehringer; (S)=Sigma; (dem.)=demaleylated

VI-I.2.2.1 Experimental details

Demaleylation of pool III (fig.VI.1) was performed in 33% acetic acid at a peptide concentration of 320 nmol/ml. The precipitate formed at addition of acetic acid remained after incubation for 69 h at 40°C. The insoluble material (Tm-7) was collected by centrifugation for 15 min at 700 x g, washed with 33% acetic acid and solubilized in 1% collidinium acetate, pH 8.3. The acid-soluble fraction was lyophilized; the dried residue was taken up into 3% pyridine + 3% picoline, pH 10.0, and the pH was adjusted to 9.0 prior to application on DEAE-Sephadex or Sephadex G-50 sf.

The acid soluble fraction of pool III (fig.VI.1) was chromatographed on a column (20 x 0.9 cm) of DEAE-Sephadex A-25 as shown in fig.VI.2. Elution was at a flow rate of 30 ml/h with 5 ml 1% pyridine + 2% collidine, brought at pH 7.3 with acetic acid, followed by a gradient going from 150 ml of this starting buffer to 150 ml 6% acetic acid; finally, the column was washed with 95% acetic acid.
For purification of Tm-1, the acid soluble fraction of pool III was first chromatographed on a column (104 x 2.3 cm) of Sephadex G-50 sf eluted with 50% acetic acid at 12 ml/h. Fractions of 3 ml were collected, of which the A280 was measured. Top fractions and tailing half of the second peak were pooled and lyophilized. This material was subjected to chromatography on DEAE-Sephadex as described above.
For purification of Tm-8, pool 2 (fig.VI.2) was fractionated on a column of Sephadex G-50 sf as shown in fig.VI.3.
The amino acid compositions of Tm-1 and Tm-7: Prolongation of the hydrolysis beyond 24 h increased the ratio numbers for Ile from 4.14 to 4.61 (Tm-1; 96 h) and from 2.95 to 3.71 (Tm-7; 72 h), because of additional bond splitting between 2 adjacent Ile residues in the peptides.

### VI-I.2.3 ISOLATION AND CHARACTERIZATION OF Tm-2 AND Tm-3 FROM POOL IV

The material in pool IV (fig.VI.1) was demaleylated and chromatographed on Sephadex G-25 sf (fig. VI.4). The main peak contained 2 peptides, Tm-2 and Tm-3. This was concluded from the amino acid composition which was not compatible with the presence of one peptide with a C-terminal Arg, and because Tyr and Ala were detected as N-terminal residues. For either peptide the results suggested a yield of about 50%.
Pool II contained a pure peptide, which in a later stage appeared to be Tm-2a, a fragment of Tm-2 comprising residue 1 to 10, in a yield of 12%. No other peptides were recognized in ninhydrin-positive fractions. The main peak was pooled, concentrated and then fractionated by paperelectrophoresis. Two major products were detected, representing

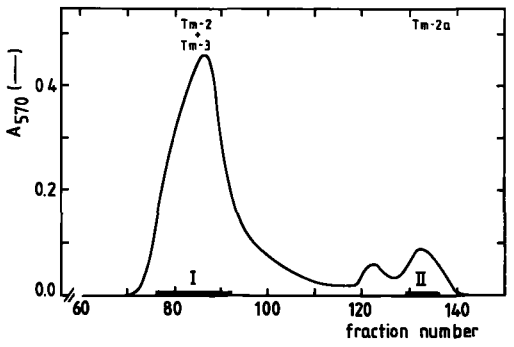



Fig.VI.4. Purification of Tm-2 and Tm-3 on Sephadex G-25 sf. The demaleylated pool IV (fig.VI.1) was fractionated on a column (128 x 2.3 cm) of Sephadex G-25 sf eluted with 5% acetic acid at 26 ml/h. Fractions of 2.6 ml were collected; aliquots (1%) were analyzed in the PA (A570:——). Fractions were pooled as indicated.

pure Tm-3 (14 residues) as the most basic spot, and practically pure Tm-2 (23 residues) a spot near the neutral amino acids; yields were 18%. The amino acid composition of both peptides is given in table VI.1. With CPA and CPB 12 residues were released from Tm-2 in quantities ranging from 0.4 to 1.0 mol per mol of peptide. Incubation of Tm-3 with LAP resulted in the liberation of all 14 residues (table VI.2). In both peptides one Asx residue was identified as Asp.

VI-I.2.3.1 Experimental details

Demaleylation of pool IV (fig.VI.1) was performed at a peptide concentration of 500 nmol/ml. The demaleylated peptide(s) were lyophilized and dissolved in 50% acetic acid prior to chromatography on Sephadex G-25 sf.

The demaleylated pool IV was chromatographed on a column (128 x 2.3 cm) of Sephadex G-25 sf eluted with 5% acetic acid at 26 ml/h (fig. VI.4). Fractions of 2.6 ml were collected; aliquots (25 µl) of fraction 60-145 were analyzed in the PA; aliquots (50 µl) of fraction 120-160 were removed for LSC. Fractions were pooled as indicated and lyophilized.

Fractionation of Tm-2 and Tm-3. The main peak, pool I, obtained on Sephadex G-25 sf was fractionated by HVE at pH 3.6 for 80 min at 4000 V. Peptides were detected on guidestrips with fluorescamine and ninhydrin II and eluted as described in chapter III.

LAP hydrolysis of Tm-3: details are given in table VI.2.

VI-I.2.4 ISOLATION AND CHARACTERIZATION OF Tm-4 FROM POOL V

Pool V contained a peptide of 7 residues, Tm-4, which was already fairly pure. Estimated yield was 69%. It was rechromatographed on Sephadex G-50 sf resulting in isolation of pure Tm-4; yield was reduced to 28% as only central fractions were pooled; the amino acid composition is given in table VI.1. CPA and CPB released all constituting residues (table VI.2).

VI-I.2.4.1 Experimental details

Purification of Tm-4 was performed on a column (108 x 2.3 cm) of Sephadex G-50 sf eluted with 50% acetic acid at 10 ml/h (result not shown). Fractions of 2.3 ml were collected, of which the A280 was measured.

VI-I.3 CONCLUSIONS

In addition to the already isolated active serine peptide (Tm-5), the 7 other tryptic fragments of maleylated atropinesterase were isolated

in a practically pure state. Their homogeneity was concluded from the results of the amino acid analyses and the endgroup determinations with dansylchloride (table VI.1). With the larger peptides (Tm-1, Tm-6, Tm-7, Tm-8), further proof of homogeneity was found in the results of a few cycles of dansyl-Edman degradation. Data obtained after digestion of some of the Tm-peptides with CPA and CPB (table VI.2) gave additional evidence; the fact that the amounts of the C-terminal residue and of some preceding amino acids liberated were equimolar to the assumed amount of peptide verified the total number of residues present.

The total number of residues of all Tm-peptides including Tm-5 (table VI.1) is in agreement with the number of residues determined in atropinesterase (table IV.3) and also the total number of residues of each amino acid agreed with the composition found for the complete protein.

## VI-II ISOLATION, PURIFICATION AND CHARACTERIZATION OF 5 CYANOGEN BROMIDE (CN-) PEPTIDES OF ATROPINESTERASE

### VI-II.1 FRACTIONATION OF THE CNBr-DIGEST

In a representative experiment 63 mg (2.1 µmol) atropinesterase was cleaved with CNBr and the peptide mixture was fractionated on Sephadex G-50 sf, with a result as is shown in fig.VI.5 (which is identical to fig.V.2). According to the elution volume and peak height, the first peak (I) probably contained a small amount of partially cleaved atropinesterase. In pool II, which contained most of the eluted material, at least 3 peptides were present, with molecular weights between 5,000 and 7,000 according to the results obtained after polyacrylamide gel-electrophoresis. Numerous attempts were made, with a multitude of approaches, to isolate these fragments, but none of the peptides could be obtained in a reasonably pure form. Fractionation was impeded in particular by the bad solubility of the CN-peptides present in pool II, the heterogeneity caused by incomplete cleavage of Met-Thr and Met-Ser bonds and the small difference in molecular weight. However, useful information about constituting fragments was obtained after redigestion of the mixture of CN-peptides ("CN-") with trypsin, pepsin

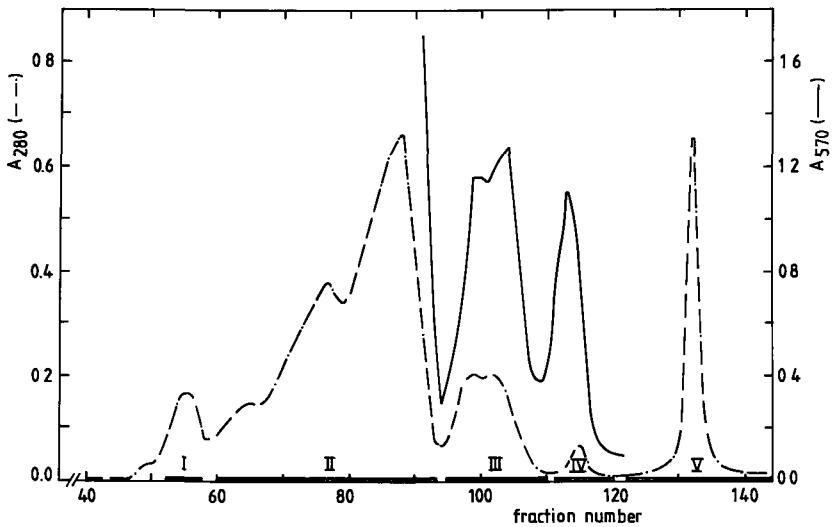



Fig.VI.5. Fractionation of CN-peptides of 2.06 µmol atropinesterase on Sephadex G-50 sf. The column (110 x 2.3 cm) was eluted with 5% acetic acid at 33 ml/h. Fractions of 3.3 ml were monitored at 280 nm (-·-·-); aliquots (1.1%) were used for the detection of ninhydrin-positive material, in the PA (A570:——). Fractions were pooled as indicated.

or S.protease, as will be presented in chapter VII.

From the remaining pools III, IV and V, after application of several purification methods, 6 pure CN-peptides were isolated including the already described active serine peptide (CN-3 = CN-a) present in pool IV, whereas a 7th fragment comprising 3 residues, was identified in impure fractions.

VI-II.1.1 Experimental details

Cyanogen bromide cleavage of atropinesterase: see chapter III.

Fractionation of CN-peptides on Sephadex G-50 sf was performed as described in V-I.1.2.1. Fractions were pooled as is indicated in fig. VI.5.

## VI-II.2 ISOLATION OF CN-PEPTIDES FROM THE DIFFERENT POOLS

### VI-II.2.1 ISOLATION AND CHARACTERIZATION OF CN-1 AND CN-8 FROM POOL III

Pool III contained a double peak (fig.VI.5) showing already some separation. Rechromatography on a longer column of Sephadex G-50 sf

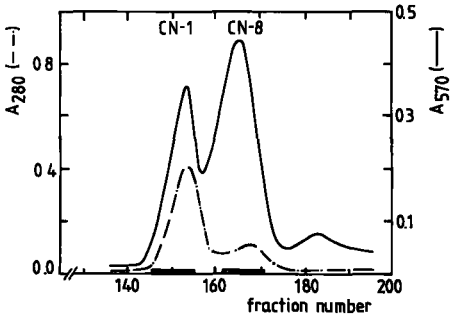



Fig.VI.6. Separation of CN-1 and CN-8 on Sephadex G-50 sf. Pool III (fig.VI.5) was fractionated on a column (280 x 1.5 cm) of Sephadex G-50 sf eluted with 5% acetic acid at 20 ml/h; fractions of 2.3 ml were collected and monitored at 280 nm (–·–·–). Aliquots (1%) were used for on line detection of peptides in the PA (A570:——). Fractions indicated by the bars contained pure CN-1 and CN-8, respectively.

yielded 2 well-identifiable peaks, as is shown in fig.VI.6. Two pools were made, each containing one pure peptide according to the amino acid analyses (ratio to Hse in relation to the approximate molecular weight) and to endgroup determinations. The first pool contained CN-1 (26 residues; see table VI.3) in a yield of 35%. This fragment was shown to be the N-terminal CN-peptide of atropinesterase because of the identical sequence of the first 5 residues; furthermore its amino acid composition agreed with the first 26 residues identified by sequence analysis of the intact protein (chapter IV). CPA only released the C-terminal Hse, while incubation with CPA and CPB liberated Ala, Ile, Lys and Hse (table VI.4).

The second pool contained CN-8 (23 residues; see table VI.3) in a yield of 38%. Dansylation did not reveal an N-terminal residue, suggesting the presence of pyroGlu. After appropriate treatment to open the ring in pyroGlu, Glu was shown to be the only N-terminal amino acid. CPA released 1 mol of Hse per mol CN-8 and 0.2 - 0.6 mol of preceding residues (table VI.4).

VI-II.2.1.1 Experimental details

Rechromatography of pool III (fig.VI.5) was performed on a column (280 x 1.5 cm) of Sephadex G-50 sf eluted with 5% acetic acid at 20 ml/h (fig.VI.6). Fractions of 2.3 ml were collected of which the A280 was measured. Aliquots (1%) of fraction 135-190 were used for peptide detection with the PA. The pools contained pure CN-1 and CN-8, respectively.

Conversion of pyroGlu to Glu was performed as described by Kawasaki and Itano (72). Freeze-dried CN-8, dissolved in 1 N HCl in dry methanol at a peptide concentration of 100 nmol/ml, was incubated at room temperature for 24 h and then lyophilized.

CPA and CPB hydrolysis: Details are given in table VI.4.

Table VI.3 Amino acid composition of cyanogen bromide (CN-)peptides of atropinesterase.

| Amino acid | CN-1 | CN-7 | CN-8 | CN-9 | CN-10 | CN-9a |
|---|---|---|---|---|---|---|
| Asx | 4.19(4) | | 3.08(3) | 1.05(1) | 2.00(2) | |
| Thr | | 0.98(1) | | | | |
| Ser | 1.94(2) | 1.01(1) | 1.01(1) | | | |
| Glx | 3.03(3) | 1.01(1) | 2.11(2) | 1.98(2) | 1.07(1) | 1.00(1) |
| Pro | 2.12(2) | | 1.07(1) | 1.08(1) | | 1.07(1) |
| Gly | | | 2.12(2) | | | |
| Ala | 4.29(4) | | 3.11(3) | | 1.10(1) | |
| Val | 1.06(1) | | | | 1.06(1) | |
| Met a) | 0.97(1) | 0.93(1) | 1.00(1) | 0.94(1) | | 0.92(1) |
| Ile b) | 4.22(5) | | 1.91(2) | | | |
| Leu | | | 1.00(1) | | 1.82(2) | |
| Tyr | | | 0.92(1) | | | |
| Phe | | | 1.85(2) | 0.96(1) | 0.97(1) | |
| His | | | 0.94(1) | | | |
| Lys | 3.17(3) | | 1.98(2) | | 2.01(2) | |
| Arg | | | 0.96(1) | | | |
| Trp | 0.73(1) | | | n.f.(1)c) | | |
| No.of residues | 26 | 4 | 23 | 7 | 10 | 3 |
| Yield | 38% | 8% | 35% | 67% | 25% | 6% |
| $NH_2$-terminus | Glu | Thr | Gln | Phe | Val | Pro |

See legends to table VI.1.
a) Met was determined as Hse.
b) The low value of Ile in CN-1 is due to the Ile-Ile bond present in CN-1.
c) No Trp was found despite the presence of 2% TGA during hydrolysis; Trp was shown to be present by specific colour reaction.

## VI-II.2.2 ISOLATION AND CHARACTERIZATION OF CN-3 AND CN-10 FROM POOL IV

The purification of these two CN-peptides had been worked out during the isolation of CN-3 (=CN-a; see V-I.1.2). Fractionation of pool IV on Sephadex G-25 gave 2 well-separated peaks as shown in fig.V.3. From the first one the fully characterized CN-3 was recovered (yield 18%). The tailing half of the second peak (pool CN-b in fig.V.3) contained a pure decapeptide, CN-10, according to amino acid analysis (table VI.3) and endgroup determination, in a yield of 25%. CN-10 should be the C-terminal fragment of atropinesterase as no Hse was present and because all constituting amino acids were identical to the residues released from Tm-8 with CPA and CPB (see table VI.4). LAP-treatment released all constituting residues (table VI.4); 1 of the 2 Asx residues was identified as Asp, the other as Asn (see table.VI.4).

Table VI.4 Amino acids released by carboxypeptidase A and B (CPA/B) or leucin-aminopeptidase (LAP) from some CN-peptides of atropinesterase.

| Peptide | Enzyme | Incubation time(h) | Amino acids released[a] |
|---|---|---|---|
| CN-1 | | | ...-Asp-Ala-Ile-Lys-Met[b] |
| | CPA/B | 18 | 0.3 0.8 0.9 1.0 |
| | " | 5 | 1.0 |
| CN-8 | | | Pro-Lys-Ala-Lys-Phe-Ile-Gln-Tyr-Asn-Gly-Phe-Gly-His-Ser-Met[b] |
| | CPA | 12 | 0.2 0.2 0.3 0.2 0.4 0.6 0.5 n.d. 0.6 1.0 |
| CN-9 | | | Phe-Trp-Glu-Asp-Pro-Glu-Met[b] |
| | CPA | 18 | 1.0 |
| | LAP | 20 | 1.0 |
| CN-9a | | | Pro-Glu-Met[b] |
| | LAP | 18 | 1.0 0.0 0.9 |
| CN-10 | | | Val-Ala-Lys-Asp-Leu-Asn-Glu-Phe-Leu-Lys |
| | LAP | 6 | 1.0 1.0 1.6 1.1 1.7 0.8 1.0 0.9 |

a) See legends to table VI.2.
b) Met was determined as Hse.

VI-II.2.2.1 Experimental details

Chromatography of pool IV (fig.VI.5) was performed as described in V-I.1.2.1.

LAP hydrolysis of CN-10 was performed as described in chapter III. Details are given in table VI.4.

## VI-II.2.3 FRACTIONATION AND CHARACTERIZATION OF CN-7, CN-9 AND CN-9a FROM POOL V

Fractionation of pool V on Sephadex G-10 (fig.VI.7) yielded 4 small and one large ninhydrin-positive peaks. The extremely high peak was identified as CNBr. Pool I contained a tetrapeptide (CN-7) and pool II a tripeptide, later identified as a C-terminal fragment of CN-9 (CN-9a), both impure. After paperelectrophoresis, CN-7 and CN-9a were obtained in a homogeneous form according to amino acid analysis (table VI.3), in yields of 8 and 6%, respectively. The low yield of CN-7, with N-terminal Thr, could be due to incomplete cleavage of a Met-Thr-bond.

Pool III contained a pure heptapeptide, CN-9, in 67% yield. According to the amino acid analysis (table VI.3) only 6 residues appeared to be present, and digestion with CPA or LAP did not indicate otherwise, as

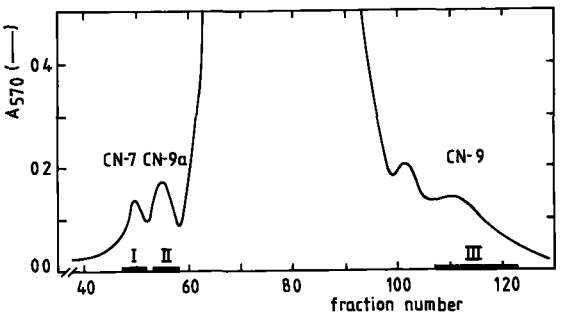



Fig.VI.7. Separation of CN-7 and CN-9 on Sephadex G-10. Pool V (fig.VI.5) was chromatographed on a column (140 x 2.3 cm) of Sephadex G-10 eluted with 5% acetic acid at 17 ml/h. Fraction size: 2.8 ml. Aliquots (1%) were used for detection of peptides in the PA (A570:——). Fractions were pooled as indicated.

only the terminal residues were released (table VI.4). However, a specific colour reaction applied to CN-9 on paper showed Trp to be present, which could not be confirmed, in various attempts, by amino acid analysis. Later sequence studies revealed the occurrence of one not identified residue next to the N-terminal Phe. This failure of identification of Trp might be due to partial destruction of this residue during cleavage with CNBr in 70% formic acid.
In the small peaks preceding pool III and in other ninhydrin-positive fractions no peptides were recognized.

VI-II.2.3.1 Experimental details
Chromatography of pool V (fig.VI.5) was performed on a column of Sephadex G-10 as shown in fig.VI.7. Fractions of 2.8 ml were collected, aliquots (1%) of fractions 41-130 were used for peptide detection with the PA.
Fractionation of pools I and II, respectively, (fig.V.7) was performed by HVE at pH 3.6 for 80 min at 4000 V. Peptides were detected on guidestrips with *fluorescamine and ninhydrin*. Staining for Trp was as described in III.1.9.2.2c. Both CN-7 and CN-9a migrated as 2 bands; this phenomenon is probably related to the presence of both Hse and Hse-lactone.

## VI-II.3 CONCLUSIONS

In addition to the already isolated active serine peptide CN-3, 5 CN-peptides of atropinesterase, and a fragment of one of them, were isolated in pure state. Their homogeneity was concluded from the amino acid compositions and from the end groups determined with dansylchloride (table VI.3). For some CN-peptides results obtained after CPA and CPB hydrolysis provided additional evidence (table VI.4).

The molecular weights calculated from the amino acid composition were in good agreement with the molecular weights based on the Kav values on Sephadex G-50 sf.
Together with CN-3, the 6 CN-peptides contained 82 of the 272 residues of atropinesterase.

## VI-III ISOLATION, PURIFICATION AND CHARACTERIZATION OF TWO PEPTIC (P-) PEPTIDES OF ATROPINESTERASE

A peptic digestion was performed with the preconceived intention to isolate 2 particular fragments, (1) a peptide connecting Tm-3 to Tm-4 in atropinesterase and (2) a peptide identical to a peptic fragment of Tm-8 (P-2; chapter VII) necessary for additional sequence studies. The peptic digest was fractionated by ion-exchange chromatography (fig. VI.8). Pools I and VII were shown to contain the desired peptides. These were purified by paper electrophoresis. From pool I pure P-B (11 residues) was obtained (over-all yield: 3%) identical to Tm-8 P-2 according to amino acid composition and endgroup determinations. From pool VII the almost pure Arg-containing fragment P-A (13 residues) was

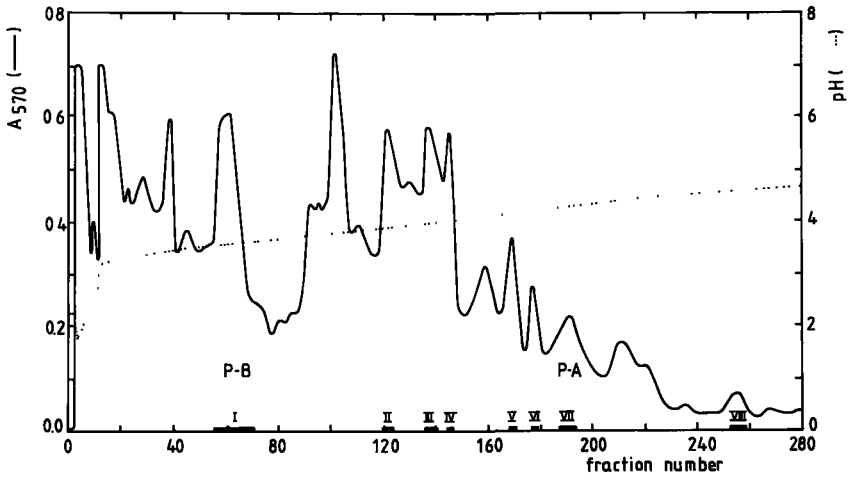



Fig.VI.8. Fractionation of P-peptides of atropinesterase. A column of Aminex A-5 (19 x 0.9 cm) was eluted at 50°C, at a flow rate of 30 ml/h. Gradient elution was applied by pumping 2.0 M pyridinium acetate, pH 5.0, into a mixing chamber containing 480 ml 0.2 M pyridinium acetate, pH 3.1, at a flow rate of 13.5 ml/h. Fractions of 1.5 ml were collected; 8% of the effluent was sacrificed for on line peptide detection in the PA (A570:——). Fractions were pooled as indicated.

Table VI.5 Amino acid composition of peptic (P-) peptides of atropinesterase.

| Amino acid | P-A | P-B |
|---|---|---|
| Asx | 1.22(1) | 1.04(1) |
| Thr | | |
| Ser | 1.09(1) | 1.07(1) |
| Glu | 0.25 | 1.07(1) |
| Pro | 1.08(1) | |
| Gly | 0.38 | 2.01(2) |
| Ala | 2.01(2) | |
| Val | | |
| Met | | 0.83(1) |
| Ile | | 0.93(1) |
| Leu | 2.87(3) | |
| Tyr | 0.84(1) | 0.97(1) |
| Phe | 0.90(1) | 1.90(2) |
| His | | 1.00(1) |
| Lys | 2.08(2) | |
| Arg | 0.92(1) | |
| No.of residues | 13 | 11 |
| Yield | 15% | (18%) 3% |
| $NH_2$-terminus | Ala | Ile |

For details see legend under table VI.1.
Yields refer to the amount of peptide isolated after one ( ) and two purification steps.

isolated. Amino acid analyses and other data are summarized in table VI.5.

From other pools more pure peptides were recovered, which yielded information (not described) that confirmed already known structures.

VI-III.1 Experimental details

Peptic digestion was performed as described in III.1.6.4 and table III.1 with 0.65 µmol atropinesterase.

The digest was fractionated on an Aminex A-5 column (19 x 0.9 cm) at 50 °C by gradient elution (see fig.VI.8). As P-A had already tentatively been identified in eluates of a peptic digest of "CN"-peptides (fig.VII-III.1), and P-B among P-peptides of Tm-8 (fig.VII-I.15), it could be guessed in what part of the eluate the products appeared. Fractions were selected on the basis of amino acid analysis. Further purification was by HVE at pH 3.6 (4000 V, 120 min).

For P-B, one of the strong His-positive, Arg-negative spots was eluted and once more purified with HVE. Homogeneity was judged by amino acid analysis and endgroup determination (table VI.5).

# CHAPTER VII THE AMINO ACID SEQUENCE OF ATROPINESTERASE

## SEQUENCE DETERMINATION OF THE Tm-PEPTIDES, OF FIVE CN-PEPTIDES AND OF FRAGMENTS CONNECTING Tm-PEPTIDES

Attempts to elucidate the amino acid sequence of all pure Tm- and CN-peptides were performed mainly by application of the dansyl-Edman degradation and sometimes with the manual Edman procedure. Of the Tm-peptides (VII-I), only the heptapeptide Tm-4 could be sequenced directly; the structure of the others, all comprising more than 10 residues, had to be elucidated after secondary cleavage(s) and sequence analysis of the smaller fragments. The whole sequence of Tm-fragments was determined or confirmed with additional evidence, obtained (1) by structural analysis of pure CN-peptides, fragments of pure CN-peptides or peptides derived from a mixture of large CN-peptides ("CN-"peptides) after secondary cleavage procedures (VII-II/III), (2) with partially structured peptic (P-)peptides of atropinesterase (VII-IV), (3) with sequence results of fragments of atropinesterase produced by uncontrolled bond splitting (VII-V). Usually, secondary cleavages were carried out with trypsin (after removal of maleylgroups from maleylated peptides), S.protease or pepsin. Sometimes chymotrypsin, thermolysin, CNBr or NBS were used. In general, tryptic digestion of peptides resulted in the production of fragments (T-peptides) which could be sequenced very well. Fragments containing internal Lys that were derived from other secondary cleavage methods were used for the ranking of T-peptides. Amides were assigned by manual Edman degradation, by the method of Offord (93) or by hydrolysis with CPA, CPB or LAP as summarized in table VII-VI.

The alignment of Tm-peptides, which resulted in the elucidation of the whole primary structure of atropinesterase, was accomplished on the basis of the results of a sequencer analysis of atropinesterase (chapter IV), of the sequence determined of one CN-peptide (CN-8), and of the sequences of some fragments of the large CN-peptides ("CN-2" p-2, sp-3a, p-4,"CN-5"p-2, sp-2a and "CN-6"p-3) (VII-III) and of one peptic peptide of atropinesterase (VII-IV), all containing internal Arg.

## VII-I SEQUENCE DETERMINATION OF Tm-PEPTIDES

### VII-I.1 Tm-1 (27 RESIDUES)

#### VII-I.1.1 SEQUENCE DETERMINATION OF Tm-1

Tm-1 was identified as the N-terminal Tm-fragment of atropinesterase because the amino acid sequence of the first 5 residues was identical to the sequence of these residues in the intact protein as determined with an automated sequencer. Furthermore, the composition of Tm-1 was in full agreement with the sequence determined for this part of the protein. However, in this sequence 2 residues, in positions 14 and 17 were not identified, so additional information remained to be found. Two tryptic fragments of Tm-1, viz. [Tm-1a] and [Tm-1b], together comprising all residues of Tm-1, which were formed during a digestion of maleylated atropinesterase by aspecific cleavage, were isolated and partially sequenced (VII-V.2). The results confirmed the sequence data and provided the missing information. Further confirmation came from the structure of CN-1 (VII-II.1), the N-terminal CN-peptide of 26 residues. The amino acid sequence of Tm-1, constructed from sequencer results of atropinesterase and sequence results of [Tm-1a], [Tm-1b] and CN-1 is shown in fig.VII-I.1. Amide assignment is summarized in table VII-VI.

### VII-I.2 Tm-2 (23 RESIDUES)

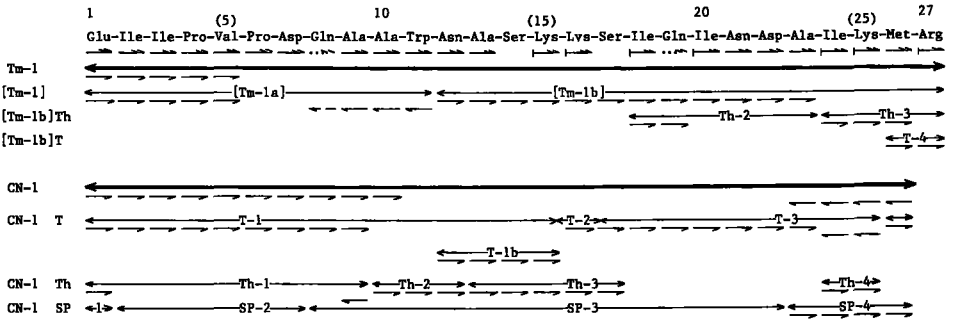
#### VII-I.2.1 SEQUENCE DETERMINATION Tm-2

Of Tm-2 the position of the first 5 residues was determined. These were identical to the first 5 residues of completely sequenced Tm-2a (see VI-I.2.3), which in this way was shown to comprise residues 1-10 of Tm-2. CPA and CPB released 1 mol Ser and 1 mol Arg per mol Tm-2.

#### VII-I.2.2 SECONDARY CLEAVAGES OF Tm-2

Treatment with pepsin resulted in the isolation of 2 peptides: P-1 comprising the first 13 residues of Tm-2, which was partially

Fig.VII-I.1. Amino acid sequence of Tm-1, constructed from sequencer results of atropinesterase, and sequence results of Tm-1, [Tm-1a], [Tm-1b], CN-1 and fragments of [Tm-1b] and CN-1.




—→ Determined by dansyl-Edman degradation.
⊢— Identified by amino acid analysis after back hydrolysis of PTH derivatives with HI.
—→ Determined as PTH amino acid by HPLC after sequencer analysis of atropinesterase.
—→ Determined as PTH amino acid by thin layer chromatography.
~→ Residues released by LAP.
←— Residues released by CPA or CPB or both.
··· Residues which were not unambiguously identified, both Gln and Glu were found on position 8 and 19 after analysis in the Beckman sequencer.
[ ] Fragments of Tm-peptides produced by uncontrolled bond splitting.

Figures refer to the positions of the residues in atropinesterase, whereas figures in parentheses refer to their positions in the peptide.
Peptides derived from tryptic cleavage of maleylated atropinesterase are designated with Tm-, peptides derived from CNBr-cleavage of atropinesterase are designated with CN-. Peptide derived from tryptic cleavage of pure Tm- and CN-peptides are designated with T-, from peptic digestion with P-, from digestion with S.protease with SP-, from thermolytic digestion with Th-, from chymotryptic digestion with Ch- and from NBS-cleavage with NBS-. Peptides derived from CNBr-cleavage of Tm-peptides are designated with CNBr-, while peptides derived from tryptic, peptic and S.protease digestion of impure "CN-"fragments are designated with t-, p- and sp-, respectively.

sequenced, and fully sequenced P-2, with the remaining 10 residues (fig.VII-I.2). Digestion of Tm-2 with S.protease yielded SP-1 (3 residues) and SP-2 (20 residues), which provided the missing sequence information.

### VII-I.2.3 SEQUENCE ASSEMBLY OF Tm-2

A survey is given in fig.VII-I.3. The N-terminal sequence of Tm-2 was extended to residue 10 by sequence analysis of Tm-2a. SP-2 extended the sequence to residue 18 and connected P-1 and P-2 by composition and partially by sequence. Fully sequenced P-2 completed the whole sequence of Tm-2. Sequence results of P-1, CPA and CPB hydrolysis results of P-1 and P-2 and LAP results of SP-1 were consistent with the established structure of Tm-2 as given in fig.VII-I.3. Assignment of amides is summarized in table VII-VI.

All sequence results of Tm-2a, Tm-2, P- and SP-peptides of Tm-2 are summarized in fig.VII-I.3.

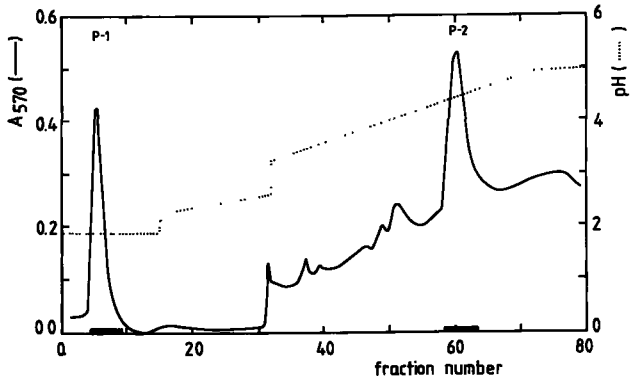



Fig.VII-I.2. Fractionation of P-peptides of Tm-2. A column of Aminex A-5 (18.5 x 0.9 cm) was eluted at 50°C, at a flow rate of 24 ml/h with 8 ml 0.05 M pyridinium acetate, pH 1.9, then 15 ml 0.05 M pyridinium acetate pH 2.4 (buffer A), followed by a two steps gradient elution. Gradients were made by pumping second buffer (B) at 10.6 ml/h into a mixing chamber with first buffer (A); subsequently third buffer (C) was pumped into the mixing chamber filled with B. Conditions: 1st gradient, 20 ml A; 2nd gradient 35 ml 0.2 M pyridinium acetate, pH 3.1 (B); buffer C, 2.0 M pyridinium acetate, pH 5.0. Elution was ended with 2 M pyridine. Fractions of 1.5 ml, representing 92% of the effluent, were collected; 8% was used for on line peptide detection in the PA (A570:——). Fractions were pooled as indicated.

VII-I.2.4 Experimental details

Digestion of Tm-2 with pepsin, fractionation and sequence determination of the P-peptides. The peptic digest of Tm-2 (III.1.7.2) was fractionated on Aminex A-5 (fig.VII-I.2). Two pure peptides, P-1 and P-2, comprising all constituting residues of Tm-2, were isolated as judged by amino acid analysis (table VII-I.1). Of P-1 the sequence of 8 amino acids was determined; CPA showed Leu to be C-terminal. According to Offord (93) the peptide had a net charge of -2, indicating the presence of 2 acidic residues and consequently 1 amide. Of P-2 the whole sequence was elucidated; Asx was identified as PTH-Asp. The DNS-derivatives of P-2 were contaminated with derivatives of the previous and the following step, starting at residue 4, partially due to incomplete coupling or cleavage during Edman degradation and partially to the presence of His at the third position (see also V-II.2.1.1). This made elucidation of the sequence Thr-Asp-Thr (6th-8th) particularly difficult. CPA and CPB showed Ser-Arg to be the C-terminal sequence (table VII-I.10).

Digestion of Tm-2 with S.protease, fractionation and sequence determination of SP-2. Fractionation of the peptide mixture obtained after digestion of Tm-2 with S.Protease (III.1.7.3) by HVE at pH 3.6 (4000 V, 120 min) gave partially pure SP-1 and pure SP-2 as judged by amino acid analysis (table VII-I.1). LAP hydrolysis of SP-1 liberated all residues in equimolar amounts (table VII-I.10), showing the presence of Glu. In SP-2, residues in positions 3 and 6-15 were identified after dansyl-Edman degradation; Asn in position 3 and Asp in position 7 in SP-2 were determined as PTH-derivatives. Identification

Table VII-I.1 Amino acid composition of P- and SP-peptides of Tm-2.

| Amino acid | P-1 | P-2 | SP-1 | SP-2 | Total Tm-2 |
|---|---|---|---|---|---|
| Asx | 1.95(2) | 1.05(1) | 0.26 | 2.91(3) | 3 |
| Thr | | 1.93(2) | | 1.92(2) | 2 |
| Ser | 1.07(1) | 1.00(1) | 0.21 | 1.91(2) | 2 |
| Glx | 1.02(1) | | 0.99(1) | | 1 |
| Pro | 2.03(2) | | 0.21 | 2.01(2) | 2 |
| Gly | 2.00(2) | 0.99(1) | 0.27 | 3.01(3) | 3 |
| Ala | 0.26 | 0.22 | | | |
| Val | 1.90(2) | | 1.07(1) | 1.10(1) | 2 |
| Met | | | | | |
| Ile | | | | | |
| Leu | 1.15(1) | 2.21(2) | 0.30 | 3.20(3) | 3 |
| Tyr | 0.82(1) | 0.75(1) | 0.93(1) | 1.00(1) | 2 |
| Phe | | | | | |
| His | | 1.12(1) | | 0.96(1) | 1 |
| Lys | | 0.34 | | | |
| Arg | | 0.95(1) | | 0.98(1) | 1 |
| Trp | 0.72(1) | | | n.d.(1) | 1 |
| No. of residues | 13 | 10 | 3 | 20 | 23 |
| Relative yield a) | 65% | 60% | 37% | 23% | |
| $NH_2$-terminus | Tyr | Leu | n.d. | n.d. | |

Amino acid analyses were performed after 18-24 h hydrolysis. Residues are reported as mol per mol peptide, according to III.1.4.2. Figures in parentheses show the number of residues of each amino acid found by sequence studies. Experimental data have been based on the results of at least duplicate amino acid analyses, unless insufficient amounts were available. Amounts taken per analysis varied from 5 to 20 nmol. Values for Thr, Ser and Trp have not been corrected for losses during hydrolysis.

a) The relative yield of the peptide refers to the amount of peptide isolated after one purification step and is relative to the amount of peptide digested; figures in parentheses (in subsequent tables) refer to yields after the final purification step.

of residues in positions 1, 2, 4 and 5 was omitted because of the low amount (25 nmol) of peptide available for sequence studies.

## VII-I.3 Tm-3 (14 RESIDUES)

### VII-I.3.1 SEQUENCE DETERMINATION OF Tm-3

The sequence of 8 residues was determined unambiguously -after modification with Braunitzer's reagent- by dansyl-Edman degradation. In addition, Ser was tentatively assigned to position 3, but was not identified with certainty because also Leu was detected, for which no explanation could be given. After incubation with LAP for 15 minutes

and 4 hours, respectively, 1.5 and 2.4 mol of Ser were liberated per mol of peptide (table VI.2). In relation to the quantity of other residues released, these values were unaccountably low and did neither prove nor disprove the presence of 2 Ser residues close to the N-terminal side or 3 Ser residues in the whole peptide. The one residue of Asx in Tm-3 was identified with LAP as Asp (tables VI.2 and VII-VI).

VII-I.3.2 SECONDARY CLEAVAGE OF Tm-3 WITH CHYMOTRYPSIN

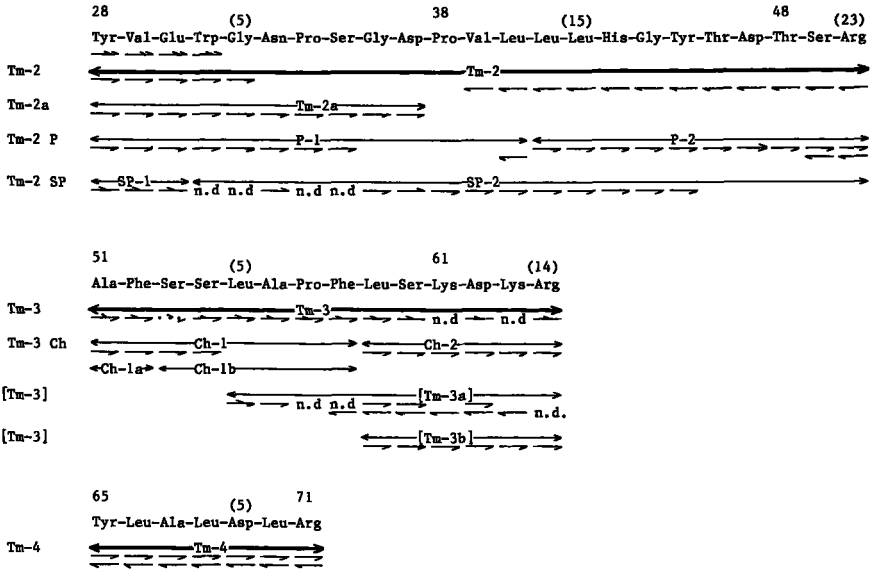
Fractionation of the digest resulted in the isolation of 4 Ch-peptides, i.e. Ch-1 and Ch-2 with the N-terminal 8 and the remaining 6 residues, respectively, and 2 fragments of Ch-1. Amino acid compositions (table VII-I.2) and sequence determination of the first 4 residues of Ch-1 gave final proof of the presence of Ser at position 3 as well as 4, and of 3 residues of Ser in total. Ch-2 was fully sequenced.

Table VII-I.2 Amino acid composition of Ch-peptides of Tm-3.

| Amino acid | Ch-1 | Ch-2 | (Ch-1a) | (Ch-1b) | Total Tm-3 |
|---|---|---|---|---|---|
| Asx | | 1.00(1) | | | 1 |
| Thr | | | | | |
| Ser | 1.95(2) | 1.02(1) | | 1.81(2) | 3 |
| Glx | | | | | |
| Pro | 1.07(1) | | | 1.22(1) | 1 |
| Gly | | | | | |
| Ala | 2.07(2) | | 1.05(1) | 1.15(1) | 2 |
| Val | | | | | |
| Met | | | | | |
| Ile | | | | | |
| Leu | 1.01(1) | 0.96(1) | | 0.92(1) | 2 |
| Tyr | | | | | |
| Phe | 1.91(2) | | 0.95(1) | 1.02(1) | 2 |
| His | | | | | |
| Lys | | 2.07(2) | | | 2 |
| Arg | | 0.96(1) | | | 1 |
| Trp | | | | | |
| No. of residues | 8 | 6 | (2) | (6) | 14 |
| Relative yield a) | 31% | 43% | 11% | 11% | |
| $NH_2$-terminus | Ala | Leu | n.d. | n.d. | |

See legends to table VII-I.1.
Peptides in parentheses contain duplicate sequences of other fragments; amino acid residues are to be ignored when the total of the fragments is computed.

Fig.VII-I.3. Amino acid sequence of Tm-2, Tm-3 and Tm-4 constructed from sequence results of Tm-2, Tm-2a, P- and SP-peptides of Tm-2, from sequence results of Tm-3 and Ch-peptides of Tm-3 and from sequence results of Tm-4.




Peptides are designated as described at fig.VII-I.1

··· Ser was not determined unambiguously in Tm-3 because also Leu was found.

## VII-I.3.3 SEQUENCE ASSEMBLY OF Tm-3

Sequence results of Tm-3, partially confirmed by the N-terminal sequence of Ch-1, combined with the sequence results of Ch-2, elucidated the whole sequence of Tm-3. Ch-1 and Ch-2 were connected by the N-terminal sequence of Tm-3, which gave an overlap of one residue. Confirmatory sequence information concerning Tm-3 was obtained from the fragments formed by aspecific cleavage of atropinesterase, viz. [Tm-3a] and [Tm-3b] described in VII-V. All sequence results are given in fig.VII-I.3.

## VII-I.3.4 Experimental details

Sequence determination of Tm-3. The peptide was modified with the Braunitzer's reagent before the first Edman cycle. Without this modification only the first 2 residues could be established because already after the second degradation cycle 60% of the degraded peptide was extracted with ethyl acetate as shown after hydrolysis of the organic phase. Extraction occurred despite the presence of polar residues such as Arg and Asp.

Chymotryptic digestion of Tm-3. The chymotryptic digest (III.1.7.5) was fractionated by HVE at pH 3.6 (4000 V, 90 min).

VII-I.4 Tm-4 (7 RESIDUES)

The sequence of all constituting residues of Tm-4 was determined unambiguously by dansyl-Edman degradation as given in fig.VII-I.3. CPA and CPB hydrolysis (table VI.2) identified Asx as Asp (table VII-VI).

VII.1.5 Tm-6 (70 RESIDUES)

VII-I.5.1 SEQUENCE DETERMINATION OF Tm-6

The sequence of the first 4 residues was determined (fig.VII-I.8), while with CPA and CPB the last 8 and 9 residues were released from the maleylated and demaleylated peptide, respectively (table VI.2).

VII-I.5.2 SECONDARY CLEAVAGES OF Tm-6

These were performed with trypsin, pepsin and S.protease, respectively, to establish the whole sequence of Tm-6; in addition, Cm-6a, a peptide identical to Tm-6 but for the absence of some C-terminal residues, was digested with trypsin. Trypsin was chosen for secondary digestion because of the relatively large number of Lys residues (7) present in Tm-6. Digestion with trypsin should produce a mixture of 8 peptides with an average length of about 9 residues, very suitable for dansyl-Edman degradation. All 8 tryptic peptides expected, with a length ranging from 3 to 15 residues and comprising all residues of Tm-6, were isolated. Six of the T-peptides were sequenced completely. The missing information on the sequence of the other 2 came from fragments of Tm-6 formed by cleavage with pepsin and S.protease. These fragments also provided part of the information needed for the alignment of the T-peptides. Final alignment was accomplished with the help of sequences of peptic or S.protease fragments of the "CN-"peptides (VII-III). The occurrence of these fragments in Tm-6 followed, a.o., from their resemblance to P- and SP-peptides derived from Tm-6.

VII-I.5.2.1 Isolation and sequence determination of T-peptides

The tryptic digests of Tm-6 and Cm-6a were fractionated by ion-exchange chromatography, as is illustrated in fig.VII-I.4.

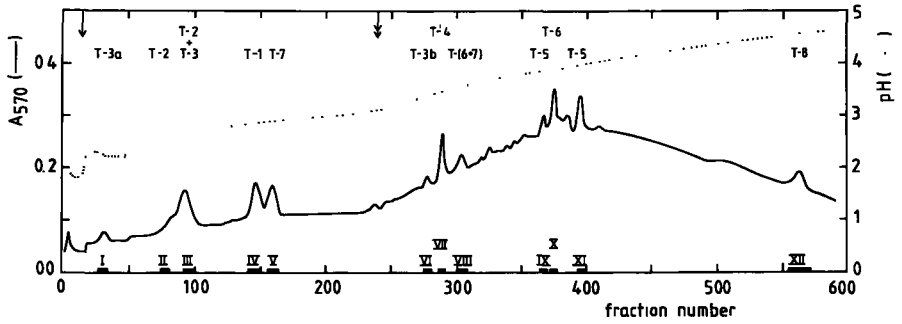



Fig.VII-I.4. Fractionation of T-peptides of Tm-6. A column of Aminex A-5 (23.5 x 0.9 cm) was eluted at 50 °C with 35 ml 0.05 M pyridinium acetate, pH 1.9, at a flow rate of 27 ml/h, followed by a two steps gradient elution (see legend to fig.VII-I.2). Buffer A: 270 ml 0.05 M pyridinium acetate, pH 2.4; buffer B: 500 ml 0.2 M pyridinium acetate, pH 3.1; buffer C: 2.0 M pyridinium acetate, pH 5.0; pumping rate: 11.7 ml/h. Between the 2 steps 25 ml B was run through the column. Elution was terminated with 42 ml 2 M pyridine: ↓ : start 1st gradient; ↓ : start 2nd gradient. Fractions of 2.0 ml were collected; 11% of the effluent was used for automatic peptide detection in the PA (A570:——). Fractions were pooled as indicated.

In the various pools, all 8 fragments except T-3 were present as practically homogeneous peptides. Tm-6 T-3 was impure, but 2 fragments of T-3, comprising all 12 residues (T-3a and T-3b), were isolated pure in fair yields. Moreover, reasonably pure T-3 was obtained from the digest of Cm-6a. Furthermore, a peptide T-(6+7), produced by incomplete cleavage at the C-terminal Lys of T-6, was isolated.

Table VII-I.3 Amino acid composition of T-peptides of Tm-6.

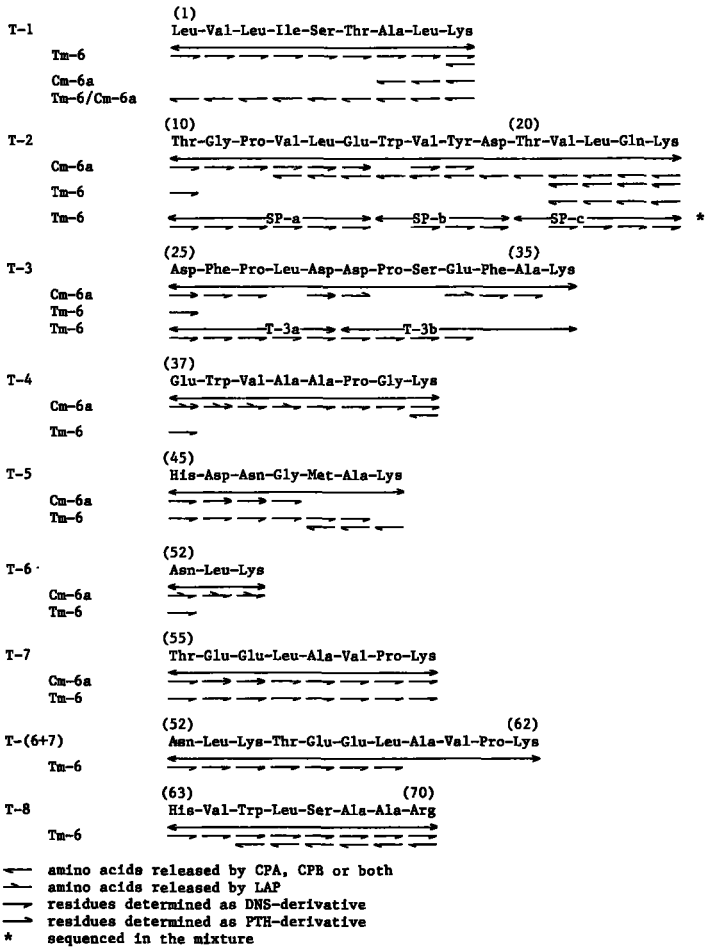
| Amino acid | T-1 | T-2 | (T-3)[b] | T-3a | T-3b | T-4 | T-5[2] | T-6 | T-7 | (T-(6+7))[b] | T-8 | Total Tm-6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | | 1.22(1) | 3.18(3) | 1.93(2) | 1.04(1) | | 1.79(2) | 1.00(1) | | 1.15(1) | | 7 |
| Thr | 0.95(1) | 1.62(2) | 0.77 | | 0.23 | | | | 0.98(1) | 1.01(1) | | 4 |
| Ser | 0.87(1) | 0.33 | 1.09(1) | 0.22 | 0.86(1) | | 0.22 | | | 0.34 | 0.98(1) | 3 |
| Glx | | 2.00(2) | 1.84(1) | 0.27 | 1.12(1) | 0.98(1) | | | 2.00(2) | 1.92(2) | | 6 |
| Pro | | 1.04(1) | 2.36(2) | 0.99(1) | 0.81(1) | 0.89(1) | | | 1.00(1) | 0.89(1) | | 5 |
| Gly | | 1.07(1) | 0.60 | 0.31 | 0.38 | 1.13(1) | 1.08(1) | 0.21 | | 0.58 | 0.21 | 3 |
| Ala | 1.09(1) | 0.25 | 1.00(1) | | 1.23(1) | 2.01(2) | 1.27(1) | | 1.06(1) | 1.25(1) | 2.03(2) | 8 |
| Val | 1.00(1) | 2.73(3) | 1.07 | | 0.29 | 1.00(1) | | | 0.94(1) | 1.11(1) | 0.91(1) | 7 |
| Met | | | | | | | 1.01(1) | | | | | 1 |
| Ile | 0.98(1) | | | | | | | | | | | 1 |
| Leu | 2.99(3) | 1.98(2) | 1.71(1) | 1.04(1) | 0.27 | | | 0.96(1) | 1.00(1) | 1.88(2) | 0.98(1) | 9 |
| Tyr | | 0.95(1) | 0.27 | | | | | | | | | 1 |
| Phe | | 0.27 | 1.91(2) | 1.04(1) | 0.85(1) | | | | | | | 2 |
| His | | | | | | | 0.95(1) | | | 0.21 | 1.11(1) | 2 |
| Lys | 0.94(1) | 1.00(1) | 1.29(1) | | 0.96(1) | 1.00(1) | 0.91(1) | 1.04(1) | 1.02(1) | 1.78(2) | | 7 |
| Arg | | | | | | | | | | | 0.99(1) | 1 |
| Trp | | 0.56(1) | | | | 0.79(1) | | | | | 0.70(1) | 3 |
| No. of residues | 9 | 15 | (12) | 5 | 7 | 8 | 7 | 3 | 8 | (11) | 8 | 70 |
| Relative yield a) | 44% | 14% | 13% | 17% | 22% | 46% | 13+21% | 40% | 46% | 15% | 23% | |
| $NH_2$-terminus | Leu | Thr | Asx | Asx | Asx | Glx | His | Asx | Thr | Asx | His | |
| Pool (fig. VII-I.4) | IV | II | III | I | VI | VII | IX+XI | X | V | VIII | XII | |

See also legends to tables VII-I.1 and 2.

b) Peptides with contaminating amino acids or peptides. In T-3a, ratio numbers deviating from integral values are due to contamination with about 35% T-2; ratio numbers were calculated by taking Ser, Ala and Phe, not present in T-2, as integral numbers.

Amino acid compositions and further results are summarized in table VII-I.3. The total number of residues in the T-peptides is in good agreement with the amino acid composition of Tm-6 (cf. table VI.1). All T-peptides derived from Tm-6 or Cm-6a except T-2 and T-3 could be fully sequenced by dansyl-Edman degradation. T-2 and T-3 were sequenced only partially; additional information came from peptides formed during digestions of Tm-6 with pepsin or S.protease. The sequences of the T-peptides and the way they have been determined are summarized in fig.VII-I.5. A detailed account is given in VII-I.5.4.1.

Fig.VII-I.5. Amino acid sequences of T-peptides of Tm-6 and Cm-6a.



VII-I.5.2.2 Isolation and sequence determination of P-peptides

Peptic digests of Tm-6 were fractionated by ion-exchange chromatography (see fig.VII-I.6). In the various pools, 4 practically pure P-peptides were present. P-6 appeared pure but for the low ratio number found for the N-terminal residue Trp. P-3 was obtained in an impure state, but could be used for sequence studies. P-4 was not isolated. Amino acid compositions and further details are summarized in table VII-I.4.

The sequence of P-3, comprising 7 residues, was completely determined, which filled-in part of the gap present in the sequence of Tm-6 T-2. Of P-6 (20 residues) the residues in 9 positions were identified, to provide for the overlap between Tm-6 T-4 and T-5. Of P-7 (13 residues) the first 7 were sequenced, to give the overlap connecting T-7 and T-8. In the other P-peptides, only the N-terminal residues were determined. A detailed account is given in VII-I.5.4.2. A survey of the sequences of the P-peptides is included in fig.VII-I.8.

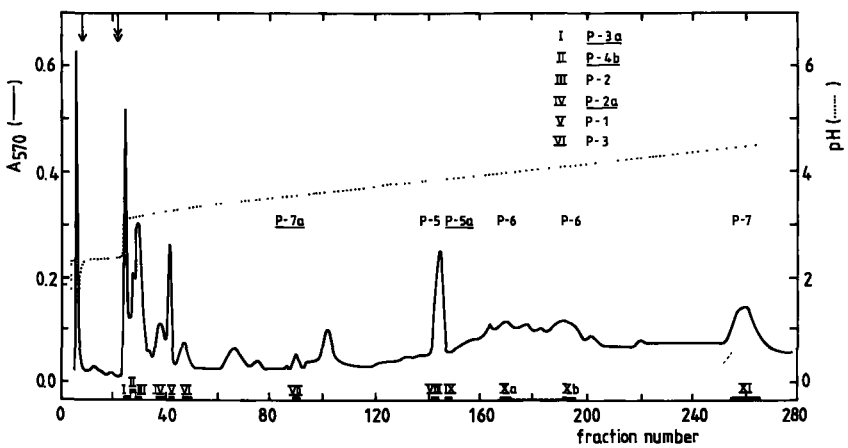



Fig.VII-I.6. Fractionation of P-peptides of Tm-6. A column of Aminex A-5 (23 x 0.9 cm) was eluted at 50 °C by two steps gradient elution (see legend to fig. VII-I.2) at a flow rate of 11 ml/h. Buffer A: 25 ml 0.05 M pyridinium acetate, pH 2.4; buffer B: 400 ml 0.2 M pyridinium acetate, pH 3.1; buffer C: 2.0 M pyridinium acetate, pH 5.0; pumping rate: 4.8 ml/h. ↓: start 1st gradient; ↓: start 2nd gradient. Fractions of 1.6 ml were collected; 15% of the effluent were used for peptide-detection in the PA (A570:——). Fractions were pooled as indicated; pools contained the peptides as mentioned, impure when underlined.

Table VII-I.4 Amino acid composition of P-peptides of Tm-6.

| Amino acid | P-1 | P-2 | P-3[b] | P-5 | P-6 | P-7 | (P-2a)[b] | (P-3a)[b] | P-4b[b] | (P-5a)[b] | (P-7a)[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | | | 1.09(1) | 0.27 | 2.73(3) | | 0.32 | 1.18(1) | 0.98(1) | 0.42 | 0.57 |
| Thr | 0.20 | 1.88(2) | 0.96(1) | | 0.90(1) | | 1.00(1) | 0.86(1) | 0.32 | | 0.24 |
| Ser | | 0.81(1) | | | 0.31 | 0.94(1) | 1.16(1) | 0.36 | 0.89(1) | | 0.26 |
| Glx | | 1.03(1) | 0.29 | 1.04(1) | 1.99(2) | | | 0.48 | 1.03(1) | 0.93(1) | 0.34 |
| Pro | | 1.02(1) | 0.28 | | 1.07(1) | 1.03(1) | | | 0.99(1) | | 1.00(1) |
| Gly | | 1.05(1) | 0.31 | 0.25 | 2.18(2) | | 0.68 | 0.36 | 0.37 | 0.35 | 0.43 |
| Ala | | 1.02(1) | 0.26 | 1.14(1) | 3.05(3) | 3.13(3) | 0.32 | 0.27 | 0.32 | 1.09(1) | 1.02(1) |
| Val | 1.16(1) | 0.97(1) | 1.83(2) | | 1.09(1) | 1.98(2) | | 1.01(1) | 0.38 | | 0.95(1) |
| Met | | | | | 0.81(1) | | | | | | |
| Ile | | 0.96(1) | | | | | 0.84(1) | | | | |
| Leu | 1.84(2) | 1.97(2) | 1.18(1) | | 1.09(1) | 1.96(2) | 0.26 | 0.95(1) | 0.28 | 0.20 | 1.03(1) |
| Tyr | | | 0.92(1) | | | | | | | | |
| Phe | | | | 0.80(1) | | | | 0.25 | | | |
| His | | | | | 0.95(1) | 1.03(1) | | | | | |
| Lys | | 0.98(1) | 0.28 | 1.02(1) | 2.96(3) | 0.98(1) | | | 0.32 | 1.00(1) | 1.22(1) |
| Arg | | | | | | 1.05(1) | | | | | |
| Trp c) | | | n.f.(0-1) | | 0.31(0-1) | 0.85(1) | | | | | |
| No. of residues | 3 | 12 | 6-7 | 4 | 19-20 | 13 | (3) | (4) | 4 | (3) | (5) |
| Relative yield a) | 43% | 30% (11%) | 51% | 14% | 10+13% | 41% | 7% (2%) | 25% (4%) | 9% | 10% | 5% |
| $NH_2$-terminus | Leu | Ile | Trp+? | Phe | Trp+? | Leu | | | | | |
| Pool (fig. VII-I.6) | V | III | VI | VIII | Xa+Xb | XI | IV | I | II | VIII | VII |
| Purification d) | | HVE | | | | | HVE | HVE | | | |

See also legends to tables VII-I.1 and 2.
b) Peptides with contaminating amino acids or peptides.
c) In P-3 and P-6, Trp was not found or in low yield despite the addition of 2% TGA. Presence was established in a different way. In the endgroup determination, next to DNS-Trp an unidentified product was found.
d) Peptides were separated on Aminex A-5, sometimes further purified by paperelectrophoresis (HVE) at pH 3.6.

VII-I.5.2.3 Isolation and sequence determination of SP-peptides

The digest resulting from the incubation of Tm-6 with S.protease was fractionated by ion-exchange chromatography (fig.VII-I.7). Five SP-peptides were isolated in reasonably pure state. Of these, SP-5 and SP-6 appeared identical to Tm-6 P-6 and P-7. One fragment (SP-3) was not obtained as a pure product. Results are summarized in table VII-I.5.

SP-2 and SP-4 (both 4 residues) were completely sequenced. Of SP-3 the first 7 (of 14) residues were positioned, providing the connection between T-2 and T-3 and filling-in the C-terminal sequence of T-2. For details see VII-I.5.4.3 and fig.VII-I.8.

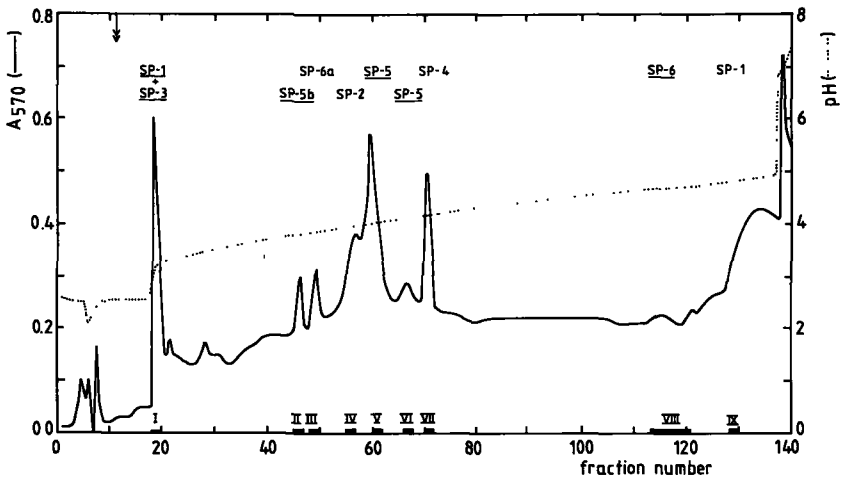



Fig.VII-I.7. Fractionation of peptides formed by digestion of Tm-6 with S.Protease. A column of Aminex A-5 (19 x 0.9 cm) was eluted at 50°C by two steps gradient elution as described for fig. VII-I.2, at a flow rate of 9.7 ml/h. Buffer A: 10 ml 0.05 M pyridinium acetate, pH 2.4; buffer B: 100 ml 0.2 M pyridinium acetate, pH 3.1; buffer C: 2.0 M pyridinium acetate, pH 5.0; pumping rate: 4.3 ml/h. Elution was terminated with 45 ml 2 M pyridine. ↓ :start 2nd gradient. Fractions of 1.1 ml were collected; 34% of the effluent was consumed by on line detection of peptides in the PA (A570: ——). Fractions were pooled as indicated; impure peptides are underlined.

Table VII-I.5 Amino acid composition of SP-peptides of Tm-6.

| | SP-1[b] | SP-2 | SP-4 | SP-5[b] | (SP-5b)[b] | SP-6[b] | (SP-6a) |
|---|---|---|---|---|---|---|---|
| Asx | 0.75 | 1.19(1) | 0.20 | 2.90(3) | 1.25(1) | | 0.34 |
| Thr | 1.83(2) | | | 1.03(1) | 0.96(1) | | |
| Ser | 0.98(1) | | | 0.72 | 0.33 | 1.21(1) | 0.20 |
| Glx | 1.34(1) | | 0.93(1) | 1.98(2) | 1.77(2) | | 0.23 |
| Pro | 1.37(1) | | | 1.12(1) | 0.22 | 1.14(1) | 0.93(1) |
| Gly | 1.03(1) | | 0.21 | 2.60(2) | 0.45 | 0.56 | 0.30 |
| Ala | 1.01(1) | | 1.11(1) | 2.97(3) | 0.38 | 2.92(3) | 1.10(1) |
| Val | 1.93(2) | 1.00(1) | | 1.33(1) | 0.25 | 1.86(2) | 1.03(1) |
| Met | | | | 0.63(1) | | | |
| Ile | 0.90(1) | | | 0.29 | | | |
| Leu | 3.57(4) | | | 1.20(1) | 0.98(1) | 1.90(2) | 0.88(1) |
| Tyr | | 0.80(1) | | | | | |
| Phe | 0.29 | | 0.79(1) | | | | |
| His | | | | 1.01(1) | | 1.06(1) | |
| Lys | 1.03(1) | | 1.17(1) | 2.73(3) | 1.04(1) | 1.01(1) | 1.04(1) |
| Arg | | | | 0.22 | | 1.06(1) | |
| Trp c) | | n.f.(1) | | n.d.(1) | | n.d.(1) | |
| No. of residues | 15 | 4 | 4 | 20 | (6) | 13 | (5) |
| Relative yield a) | 6% | 12% | 66% | 10+4% | 17% | 8% | 28% |
| $NH_2$-terminus | | Trp | Phe | | Asx | | Leu |
| Pool d) | IX | IV | VII | V+VI | II | VIII | III |

See legends to tables VII-I.1 and 2.
b) Peptides with contaminating amino acids or peptides.
c) Trp was not found despite the addition of 2% TGA or not determined.
d) Peptides were separated on Aminex A-5; see fig.VII-I.7.

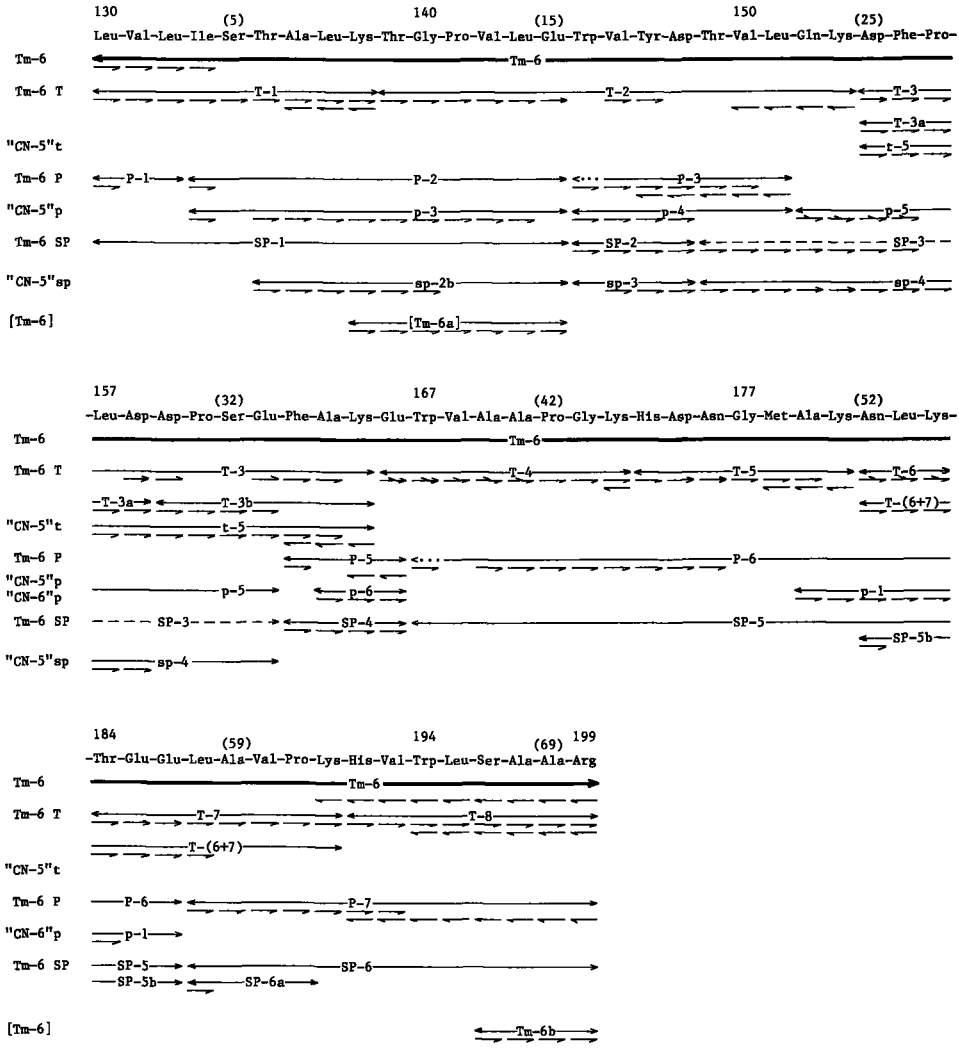
VII-I.5.3 SEQUENCE ASSEMBLY OF Tm-6

A structure for the 70 residues comprising tryptic peptide Tm-6 could be constructed from the T-peptides of Cm-6a and Tm-6 and from peptides overlapping these T-peptides by amino acid composition and/or sequence obtained after digestion of Tm-6 with pepsin or S.protease, or after digestion of "CN-"peptides with the same enzymes (VII-III). This assembly is surveyed in fig.VII-I.8. Six of the 8 T-peptides were completely structured by direct sequence analysis. T-2 needed additional data, which came from partially or fully structured Tm-6 P-3, Tm-6 SP-2 and SP-3 and "CN-5"sp-4. The sequence of T-3 was completed and confirmed by the sequences of "CN-5"t-5 and Tm-6 SP-4.

T-1 and T-2 were connected Tm-6 P-2 and SP-1 by composition and by "CN-5"p-3, a peptide identical to Tm-6 P-2, and "CN-5"sp-2b by sequence. The connection between T-2 and T-3 was provided by the identical peptides Tm-6 SP-3 and "CN-5"sp-4, the sequence data of which overlapped T-2 by 5, and T-3 by 2 and 5 residues, respectively. T-3 and T-4 (only one residue) were overlapped by Tm-6 P-5 and the identical, fully sequenced Tm-6 SP-4 and by "CN-5"p-6. Nevertheless T-4 was localized unambiguously, as it was the only T-peptide with N-terminal Glu. The identical peptides Tm-6 P-6 and Tm-6 SP-5, comprising 20 residues, overlapped T-4, T-5, T-6 and T-7 by composition, while sequence data of Tm-6 P-6 connected T-4 to T-5. T-5 and T-6 were overlapped by the partially sequenced fragment "CN-6"p-1, which also overlapped the first residue of T-7. The link between T-6 and T-7 was confirmed by the sequence of T-(6+7), which overlapped T-7 by 4 residues. T-7 and T-8 were connected by composition of the identical C-terminal fragments Tm-6 P-7 and SP-6 and by sequence analysis of P-7. CPA and CPB hydrolysis of intact Tm-6 and of Tm-6 P-7 confirmed the correct positioning of P-7 at the C-terminus of Tm-6.

The amino acid composition and/or sequence results of other P- and SP-peptides derived from Tm-6 or from the "CN-"peptides confirmed the ranking of the T-peptides mentioned above. The characterization of all amides present in Tm-6 as determined in T-peptides of Tm-6 and Cm-6a (VII-I.5.2.1), in P- and SP-peptides of Tm-6 (VII-I.5.2.2/3) and in fragments of "CN-"peptides (VII-III) is summarized in table VII-VI.

Fig.VII-I.8. Amino acid sequence of Tm-6 constructed from T-, SP and P-peptides of Tm-6 and p-, sp- and t-peptides of "CN-"fragments.




Peptides are designated as described at fig.VII-I.1

- - - SP-3 was not sufficiently pure for accurate determination of amino acid composition; sequencing was possible, though.

. . . In P-3 and P-6 Trp was not found or in low yield after amino acid analyses; in the endgroup determination, next to DNS-Trp an unidentified product was found.

VII-I.5.4 Experimental details

VII-I.5.4.1a Tryptic digestion of Tm-6 and Cm-6a and fractionation of tryptic (T-) peptides. Demaleylated Tm-6 (400 nmol) and Cm-6a (550 nmol) were digested with trypsin as described in chapter III. The digests were fractionated on Aminex A-5. In the first experiments T-peptides of Cm-6a were eluted with 25 ml 0.05 M pyridinium acetate, pH 2.4, followed by 25 ml 0.2 M pyridinium acetate, pH 3.1, and one gradient going from 0.2 M pyridinium acetate, pH 3.1, (400 ml) to 2 M pyridinium acetate, pH 5.0 (elution profile not shown). This resulted in a poor resolution below pH 3.0. Separation was markedly improved when a two-steps gradient in pH and pyridine concentration was used, as shown for the fractionation of Tm-6 T-peptides in fig.VII-I.4.

VII-I.5.4.1b Characterization of T-peptides of Tm-6. The pools obtained after fractionation of the T-peptides of Tm-6 on Aminex A-5 contained 7 peptides, Tm-6 T-1, T-2, T-4, T-5, T-6, T-7 and T-8 in a homogeneous form as judged by amino acid analysis and endgroup determination with dansylchloride. T-5 was eluted at 2 different pH values (pH 3.82 and 3.95); the peptide eluted at the lower pH, T-$5^1$, contained some impurities; T-$5^2$, eluted at pH 3.95, was isolated in pure state. Only pool III appeared to consist of 2 T-peptides: beside T-3 about 35% T-2 was present according to the amino acid analysis (the amount and the amino acid composition of T-3 could be calculated by correcting for the known amino acid composition of T-2 and from the amounts of Ser, Ala and Phe which did not occur in T-2; the composition of T-3 thus obtained agreed with the residues present in T-3a + T-3b). Beside the peptides expected after digestion of Tm-6 with trypsin, some fragments of T-peptides and one peptide containing internal Lys were isolated. Beside the impure peptide T-3, pure T-3a and T-3b were found as a result of partial cleavage between Asp 29 and Asp 30. Fractions between pools X and XI contained a mixture of T-5b and T-8b as judged by amino acid analysis and reaction with dansylchloride. T-5b, comprising the last 5 residues of T-5, was produced by partial cleavage behind Asp 46, relative yield being 5%. The production of T-3a, T-3b and T-5b probably had occurred during removal of maleylgroups in 33% acetic acid, because bond breakage between 2 Asp residues can occur in an acidic medium. Presumably, Asx 47 originally was the amide, but bond breakage between Asp 46 and Asn 47 became possible after partial conversion of Asn to Asp, the occurrence of which was deduced during sequence determination of Cm-6a T-$5^1$ and T-$5^2$ (VII-I.5.4.1d). This partial conversion also explained the elution of T-5 from Aminex at 2 different positions. T-8b, comprising the last 4 residues of T-8, was produced by partial hydrolysis behind Leu 66, relative yield being 7%. Easy rupture of this bond was also shown by isolation of Cm-6a (VI-I.2.1). No attempt was made for further purification or characterization of T-5b and T-8b. Incomplete cleavage at Lys 54 produced pure T-(6+7) beside the peptides T-6 and T-7, possibly because of incomplete removal of maleylgroups. Aspecific bond breakage due to chymotryptic-like action had not occurred to a detectable extent.

For all the mentioned peptides except T-3, T-5b and T-8b, well-fitting amino acid compositions were found. Relative yields were 14 to 46%. For details of all T-peptides see table VII-I.3. The sum of the amino acid compositions of all Tm-6 T-peptides, except T-3, but including the aspecifically formed fragments T-3a and T-3b was in good agreement with the amino acid composition as found for Tm-6 (table VI.1).

VII-I.5.4.1c Characterization of T-peptides of Cm-6a. Cm-6a T-1, T-2 and T-3 were isolated in a heterogeneous form. According to the amino acid analysis Cm-6a T-1 was contaminated with T-7, T-2 was contaminated with T-3 (an impurity also found with CPA and CPB (table VII-I.11) and by sequence analysis of the impure peptide) and T-3 was slightly contaminated with T-2. Impurities did not disturb sequence determination (VII-I.5.4.1d). The amounts were too small for further purification. T-4, T-5, T-6 and T-7 were isolated in pure state, according to the amino acid analysis. T-5 was eluted at 2 different pH values (3.83 and 3.97) as was found for Tm-6 T-5. The peptide eluted at pH 3.83, T-5[1], was homogeneous. T-5[2], eluted at the higher pH, was obtained in pure state after HVE at pH 6.5 (4000 V, 30 min). T-1, T-2, T-3 and T-7 were eluted at higher pH values than the corresponding T-peptides of Tm-6 due to other conditions during ion-exchange chromatography. A peptide corresponding with Tm-6 T-8 was not found. An impure fragment containing the first 4 residues of T-8 was tentatively identified at about pH 4.5 (relative yield 4%). The amino acid compositions of the pure T-peptides of Cm-6a are not shown because they were identical to the compositions of the corresponding T-peptides of Tm-6 (table VII-I.3).

VII-I.5.4.1d Sequence determination of T-peptides of Tm-6 and Cm-6a. Routinely the dansyl-Edman procedure was used. Manual Edman degradation was performed to distinguish between acidic residues and their amides. All sequence results are given in figs.VII-I.5 and 8. After incubation of T-peptides with CPA/CPB and LAP, C-terminal and N-terminal residues, respectively, were identified, including acidic amino acids as shown in fig.VII-I.5 and quantitatively in table VII-I.11.
T-1: Of Tm-6 T-1 the sequence of all 9 residues was determined. Val in position 2 gave 2 DNS-derivatives, being DNS-Val and DNS-Val-Leu (III.1.10.1b). C-terminal Lys was confirmed with CPB. After incubation of Cm-6a T-1 with CPA and CPB for 10 minutes, the last 3 residues were released, confirming the C-terminal residues. The quantity of each amino acid released per mol Cm-6a T-1 and Tm-6 T-1 after longer incubations confirmed the amino acid composition. The liberation of C-terminal residues by CPA seemed to be retarded by the presence of Thr at the C-terminal side.
T-2: Of Tm-6 T-2 the first residue was identified with dansylchloride; of Cm-6a T-2 the N-terminal sequence of 6 residues and the amino acids in positions 8 and 9 were determined with certainty, despite contamination with T-3. Glx in position 6 was identified as PTH-Glu. In position 7 no residue was found probably due to destruction of Trp. Beside DNS-Val in position 8 a second DNS-derivative was found with the Rf of DNS-bis-Tyr, apparently derived from DNS-Val-Tyr (III.1.10.1b). After 21 h incubation of Cm-6a T-2 with CPA and CPB, all residues up to Pro

in position 3 were released, including 1 Asp, 1 Trp, 1 Gln and 1 Glu. The detection of Glu is consistent with the identification of PTH-Glu in position 6 during sequence analysis. The high content of "Thr" is due to the presence of Gln beside Thr. After shorter incubations with CPA and CPB, 1 mol Lys and about 0.5 mol of the preceding residues were released from Cm-6a T-2 and Tm-6 T-2. (Excess amino acids liberated from Cm-6a T-2 could be ascribed to contaminating T-3.) The amino acid sequence of T-2 was completely established with P- and SP-peptides from Tm-6 (VII-I.5.4.2/3) and with p- and sp-peptides of "CN"-fragments (VII-III) as described in VII-I.5.3 and shown in fig.VII-I.8. Sequence results obtained after digestion of Tm-6 T-2 with S.protease (III.1.7.3) and Edman degradation of the SP-peptide mixture were consistent with the sequence given. The 3 SP-peptides (SP-a,b,c) were sequenced simultaneously; the small amount of peptide material available (30 nmol) made fractionation impossible. After 5 subsequent degradation cycles DNS-derivatives corresponding with the N-terminal sequences of Tm-6 T-2, of Tm-6 SP-2 (= "CN-5" sp-3) and of Tm-6 SP-3 (= "CN-5" sp-4) were identified simultaneously. Furthermore, these results confirmed that Tm-6 T-2 comprised Tm-6 SP-2, Tm-6 P-3 and part of Tm-6 SP-3.

T-3, T-3a and T-3b: Of the pure peptides T-3a and T-3b, the sequence of 5 and 4 residues respectively was determined. Of Tm-6 T-3 only the first residue was identified, while dansyl-Edman and manual Edman degradation of Cm-6 T-3 confirmed the sequences in T-3a + T-3b with regard to the positions 1 to 3, 5, 6 and 9 (in T-3), and extended the sequence to Ala 11, despite the presence of T-2. Asx residues in positions 1, 5 and 6 and Glx in position 9 were identified as PTH-Asp and PTH-Glu. According to Offord the peptide had a net charge of -3, being in accordance with the presence of 4 acidic and 1 basic amino acid. The determined structure of T-3 was confirmed and completely established with additional evidence obtained from an identical, fully sequenced peptide "CN-5"t-5 (VII-III) as shown in fig.VII-I.8.

T-4: Of Cm-6a T-4 the sequence of all 8 residues was determined; the first 2 residues Glu and Trp were also identified as their PTH-derivatives. With LAP the presence of Glu and Trp in the N-terminal part of the peptide was confirmed. With CPA and CPB only Lys and a small amount of Gly were released due to Pro in position 6 (Pro partially or completely prevents liberation of the residue at its C-terminal side). Of Tm-6 T-4 only the N-terminal residue was identified.

T-5: Of Tm-6 T-$5^2$ the sequence of 6 residues was determined; with CPA and CPB 3 C-terminal residues were found, giving the whole sequence. Of both peptides Cm-6a T-$5^1$ and T-$5^2$ the sequence of 4 residues was established with dansyl- and manual Edman degradation. In both peptides PTH-Asp was found at position 2, at position 3 PTH-Asp was found in T-$5^1$ and PTH-Asn in T-$5^2$.

T-6: The first residue of Tm-6 T-6 was identified and the complete sequence of Cm-6a T-6. With LAP all 3 residues were released, showing the presence of Asn.

T-7: Both peptides Cm-6a T-7 and Tm-6 T-7 were completely sequenced. Residues in positions 2 and 3 were identified as PTH-Glu in Cm-6a T-7.

T-(6+7): The amino acid sequence of the first 7 residues was determined. The N-terminal sequence together with the amino acid composition gave the arrangement of the tryptic peptides T-6 and T-7.
T-8: This peptide was completely structured, while with CPA and CPB 6 residues were released. One peptide of Cm-6a, being tentatively identified as the N-terminal part of T-8, was definitely characterized by amino acid sequence results of the first 2 residues, which were identical to the first 2 residues of T-8.
As shown in fig.VII-I.5, almost all residues in Tm-6 were unambiguously positioned in one of the tryptic peptides of Tm-6 or Cm-6a by direct sequence analysis. Only sequence elucidation of T-2 and T-3 needed the additional evidence mentioned above.

VII-I.5.4.2a Digestion of Tm-6 with pepsin, fractionation and characterization of peptic (P-)peptides. Demaleylated Tm-6 was digested with pepsin as described in III.1.7.2. After fractionation of the peptide mixture on Aminex A-5 (fig.VII-I.6), P-1, P-2, P-5 and P-7 were isolated in a homogeneous form as judged by amino acid analysis. In pool Xa P-6 was found with a well-fitting amino acid composition, only the stoichiometric value for Trp was low. The same peptide was found in pool Xb and in fractions between these pools. P-3 was recognized in pool VI with some impurities. This peptide did not contain Trp according to the amino acid analysis. However, dansylation showed Trp to be the N-terminal residue of both P-3 and P-6. P-4b, P-5a and P-7a were isolated in impure fractions of pool II, IX and VII, respectively. After HVE of pools I and IV at pH 3.6 (4000 V, 120 min), P-3a and P-2a were recognized in impure eluates. These fragments did not provide any new information. No fragments were recognized in other parts of the eluate. Amino acid composition, relative yields and N-terminal residues of pure P-peptides of Tm-6 indicated in fig.VII-I.6 are given in table VII-I.4.
Characterization of P-6 and P-3 was difficult because P-6 was spread out over many fractions and in both peptides no Trp, or only a small amount, was found. The spread-out of P-6 might be explained by the large number of residues, while the low value for Trp could be due to destruction during hydrolysis, or these phenomena might be explained by the specific action of pepsin, an enzyme which is able to split peptide bonds on either side of aromatic residues producing a mixture of 2 almost identical fragments. One of these fragments is foreshortened at the N- and/or C-terminal side by one residue compared to the other. In case of P-3 and P-6 partial hydrolysis might have occurred at the N-terminal side, resulting in a mixture of a peptide with N-terminal Trp and an almost identical peptide without Trp. It might be difficult to fractionate such a mixture, which should give a well-fitting amino acid composition with only a low value for Trp. The spreading of apparently homogeneous peptide material was also encountered during isolation of the peptic peptide Tm-5 P-4 (chapter V-II.2.2.3), where the presence of 2 peptides differing only by the N-terminal Phe was revealed.

VII-I.5.4.2b Sequence determination of P-peptides of Tm-6. Sequence determination of Tm-6 P-peptides was performed in a similar way as described for Tm-6 T-peptides.

Tm-6 P-1, P-2 and P-5: Only the first residues of P-1, P-2 and P-5, which confirmed parts of Tm-6 by composition, were determined with dansylchloride because of the analogy with already known sequences of Tm-6 SP-peptides and "CN-5"p- and sp-peptides. From P-5 2 residues were released with CPA in equimolar amounts confirming the presence of Glu in position 37 in Tm-6.

Tm-6 P-3: Sequence analysis of P-3 was performed to determine an internal sequence of Tm-6 T-2 and to determine the joint between Tm-6 SP-2 and SP-3. During sequence determination the same problems were encountered as described for P-6 (see below), however less obvious. The unknown DNS-derivative in position 1 and the "contamination" of a DNS-derivative with a subsequent residue, taking Trp as the N-terminal residue, suggested that the same phenomenon had occurred at Trp 16 in Tm-6 as at Trp 38. Nevertheless the sequence of 6 residues of P-3 could be established. Five residues were liberated with CPA which were consistent with the C-terminal sequence of Tm-6 SP-2 and the N-terminal residues of SP-3. The presence of Asp at position 19 in Tm-6 was confirmed.

Tm-6 P-6: the sequence of 9 residues was determined unambiguously. The presence of Trp at the N-terminal side was shown by dansylation. The established sequence gave the desired overlap from T-4 to T-5. It has to be noted that beside Trp in position 1 a second, unknown DNS-derivative was found, at the Rf of DNS-Phe. This product did not disappear after hydrolysis for 20 h, while after each Edman degradation cycle 2 DNS-derivatives were detected, one being in accordance with Trp as N-terminal residue, the second one corresponding with the residue one Edman cycle ahead. Two assumptions were made for this occurrence: the peptide material existed of 2 almost identical peptides differing only by the N-terminal residue (VII-I.5.4.2). This explanation accounted for the low value for Trp but not for the unknown DNS-derivative at position 1 or for the absence of the first DNS-derivative of the foreshortened peptide. A second assumption holds that Trp might have been partially converted into a derivative during incubation with pepsin in 5% formic acid, resulting in the presence of an unknown derivative after dansylation. Release of this derivative should have occurred simultaneously with the second amino acid to explain the detection of the "third" residue in position "2" as well as the detection of the next residues one cycle ahead, taking Trp as the first residue. The possibility of such an event is known to occur when His is encountered during sequence analysis (chapter V).
Because the sequence to be established could be determined without uncertainty no further investigations were undertaken to clarify this phenomenon.

Tm-6 P-7: The sequence of 7 residues of P-7 was determined giving the desired overlap from T-7 to T-8. With CPA and CPB 8 residues were released, which fitted in the C-terminal sequence determined for T-8.
The established sequences of P-peptides of Tm-6 are given in fig.VII-I.8. Results from CPA and CPB incubations are shown in table VII-I.11.

VII-I.5.4.3a Digestion of Tm-6 with S.protease, fractionation and characterization of S.protease (SP-)peptides. Demaleylated Tm-6 was digested with S.protease as described in III.1.7.3. 200 nmol of the peptide mixture were acidified to a final pH of about 1 and fractionated on Aminex A-5 (fig.VII-I.7). A mixture of Tm-6 SP-1 and SP-3 was identified in pool I according to the amino acid analysis and the dansyl-Edman degradation, relative yields being 20% and 30% respectively. Remarkably, SP-1 was also eluted in pool IX at pH 4.7 in an almost homogeneous form. In SP-2 (pool IV) 3 residues were identified by amino acid analysis. By dansylation, however, a 4th residue, Trp, was detected as the N-terminal residue. SP-4, SP-5 (which was identified in 2 different pools) and SP-6 were isolated in pure state or with some contaminating amino acids. SP-5 and SP-6 were identical to Tm-6 P-6 and P-7, respectively, according to amino acid composition and they also were eluted at comparable pH values. Beside the mentioned SP-peptides which comprised all residues present in Tm-6, 2 fragments of SP-peptides were isolated: SP-5b, comprising residues 52-57 of Tm-6, and SP-6a, identical to P-7a, comprising residues 58-62. SP-5b still contained some contaminating residues. These peptides were produced by partial aspecific hydrolysis behind Lys 51 and Lys 62, respectively. Incomplete removal of maleylgroups from Lys could have made these residues susceptible to hydrolysis by S.protease due to their acidic character. In other parts of the eluate no distinct fragments were recognized.
Amino acid compositions, relative yields and N-terminal residues of pure or almost pure SP-peptides of Tm-6 are given in tabel VII-I.5.

VII-I.5.4.3b Sequence determination of SP-peptides of Tm-6.
Tm-6 SP-2: The sequence of SP-2 was determined, definitely showing the presence of Trp at position 16 and Asx at position 19 in Tm-6. Val at position 2 gave beside DNS-Val a second product which should be derived from DNS-Val-Tyr as was found during sequence analysis of Cm-6a T-2. The established sequence was identical to the one of "CN-5"sp-3 (VII-III).
Tm-6 SP-3: The N-terminal sequence of 7 residues of SP-3 could be determined unambiguously in the peptide mixture of SP-1 and SP-3, because the N-terminal sequence of SP-1 was known from analysis of Tm-6 T-1. The sequence results of SP-3, giving the C-terminal sequence of Tm-6 T-2 and connecting T-2 and T-3, were confirmed by structural analysis of the pure identical peptide "CN-5"sp-4 (VII-III).
Tm-6 SP-4: Sequence analysis of SP-4, a peptide identical to Tm-6 P-5 according to the composition and the partially determined sequence, established the order of T-3 and T-4.
Tm-6 SP-1, SP-5 and SP-6: SP-1, overlapping Tm-6 T-1 and T-2 by composition, was not sequenced because the necessary information to align T-1 and T-2 was already available in "CN-5"p-3 and sp-2b (VII-III). SP-5 and SP-6 were not sequenced because of the partially known sequences of the identical peptides P-6 and P-7, giving the connections between T-4 and T-5 and between T-7 and T-8.
Tm-6 SP-5b and SP-6a: Only the first residue of these peptides was identified.

The specific action of S.protease confirmed the acidic character of Glu 15, Asp 19, Glu 33, Glu 37 and Glu 57 in Tm-6.
For the established sequences of SP-peptides of Tm-6 see fig.VII-I.8.

## VII-I.6 Tm-7 (38 RESIDUES)

### VII-I.6.1 SEQUENCE DETERMINATION OF Tm-7

The identity of 13 residues was determined -after modification of free aminogroups with Braunitzer's reagent- with the dansyl-Edman procedure. The 6th residue was identified as PTH-Asn. The 7th position was tentatively assigned to Trp.

### VII-I.6.2 SECONDARY CLEAVAGES OF Tm-7

Digestions were performed with trypsin, S.protease, thermolysin, CNBr and NBS, respectively. One SP-peptide was digested again with chymotrypsin. Of each peptide mixture one or more fragments were necessary to establish the complete sequence of Tm-7. Internal sequences were confirmed with some thermolytic peptides. All constituting residues of Tm-7 were recognized in T-, SP- and CNBr-peptides.

#### VII-I.6.2.1 Isolation and sequence determination of T-peptides

Fractionation of the tryptic digest on Sephadex G-15 yielded a peak of Tm-7 T-2 (fig.VII-I.9), a pentapeptide that was sequenced all through. Isolation of the other 2 T-peptides was hampered by the strong tendency to form aggregates. Reasonably pure T-3 was isolated after rechromatography on Sephadex G-25 sf (fig.VII-I.10). Pure T-1 was obtained

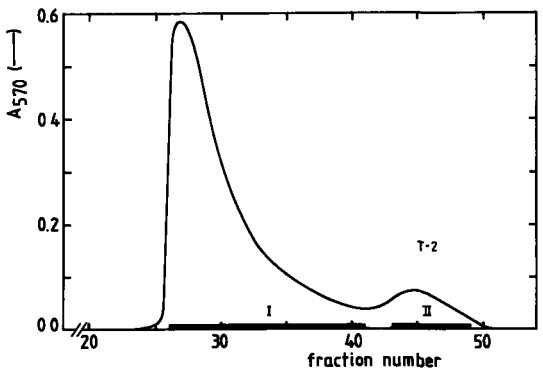



Fig.VII-I.9. Fractionation of T-peptides of Tm-7. A column of Sephadex G-15 (62 x 2.0 cm) was eluted with 1% collidinium acetate, pH 8.3, at 18.8 ml/h. Fractions of 2.5 ml were collected. Aliquots (2%) were analyzed with the PA (A570:——). Pool II contained Tm-7 T-2.

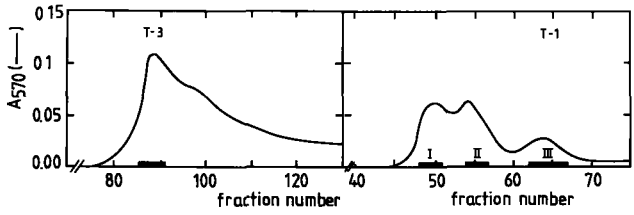



Fig.VII-I.10. (left figure) Refractionation of T-peptides of Tm-7. Pool I obtained after gelfiltration on Sephadex G-15 of the tryptic digest Tm-7 (fig.VII-I.9) was fractionated on a column of Sephadex G-25 sf (120 x 2.3 cm), eluted with 5% acetic acid at 29 ml/h. Fractions of 2.2 ml were collected. Peptides were detected with the PA (A570:——). The pool contained Tm-7 T-3.
Fig.VII-I.11. (right figure) Fractionation of T-peptides of Tm-7. A column of Sephadex G-25 sf was eluted with 33% acetic acid at 9 ml/h. Fractions of 1.0 ml were collected; 35% of the effluent was used for on line detection of peptides in the PA (A570:——). Pool III contained Tm-7 T-1.

by gelfiltration of the tryptic digest in 33% acetic acid (fig.VII-I.11), but in all cases yields were low. Results are summarized in table VII-I.6. The sequence of the first 3 residues and the amino acid composition of T-1 agreed with the known N-terminal sequence of Tm-7.

Table VII-I.6 Amino acid composition of peptides obtained after digestion of Tm-7 with trypsin and thermolysin, respectively.

| Amino acid | T-1 | T-2 | T-3 | Total Tm-7 | Th-1 | Th-2 | Th-3 | Th-4 |
|---|---|---|---|---|---|---|---|---|
| Asx | 1.23(1) | | 3.73(4) | 5 | | | 1.91(2) | |
| Thr | 1.06(1) | 1.15(1) | 1.90(2) | 4 | | | 0.79(1) | |
| Ser | 2.05(2) | 0.30 | 1.03(1) | 3 | 1.10(1) | 0.93(1) | 1.06(1) | |
| Glx | | | 3.95(4) | 4 | | | 3.04(3) | |
| Pro | | | 1.15(1) | 1 | | | 1.03(1) | |
| Gly | 0.91(1) | 0.21 | 1.29(1) | 2 | 1.04(1) | | 1.25(1) | |
| Ala | 1,86(2) | 1.12(1) | 0.35 | 3 | | 1.12(1) | | |
| Val | | | | - | | | | |
| Met | | | 2.18(2) | 2 | | | 0.70(1) | |
| Ile b) | 1.03(2) | | 1.95(2) | 4 | | | 0.55(1) | 0.96(1) |
| Leu b) | | 0.99(1) | 2.17(2) | 3 | | | 1.22(2) | |
| Tyr | | 0.73(1) | | 1 | | 0.71(1) | | |
| Phe | 0.88(1) | | | 1 | 0.86(1) | | | |
| His | | | | - | | | | |
| Lys | 0.97(1) | 1.02(1) | | 2 | | 1.24(1) | | |
| Arg | | | 0.91(1) | 1 | | | | 1.04(1) |
| Trp | n.d.(1) | | 0.82(1) | 2 | | | n.d.(1) | |
| No. of residues | 12 | 5 | 21 | 38 | 3 | 4 | 14 | 2 |
| Relative yield a) | 27% | 22% | (17%) | | 15% | 15% | 15% | 50% |
| $NH_2$-terminus | Gly | Tyr | Thr | | Gly | Ala | Leu | Ile |
| Purification c) | G-25 33% acetic acid | G-15 | G-15 G-25 5% acetic acid | | HVE | HVE | HVE | HVE |

See also legends to table VII-I.1
b) Low stoichiometric values of Ile in T-1 and Ile and Leu in Th-3 were due to incompletely hydrolyzed peptide bonds (e.g. Ile-Ile).
c) Peptides were purified by gelfiltration on Sephadex G-25 sf (G-25) or G-15 (G-15), or by paperelectrophoresis at pH 3.6 (HVE), as indicated.

Of Arg-containing T-3, the first 8 residues were sequenced. Sequence results are included in the survey of fig.VII-I.13. For details see VII-II.6.4.1.

VII-I.6.2.2 Isolation and sequence determination of SP-peptides

The digestion resulted in an acid-soluble and acid-insoluble fraction. From the former, 3 small peptides were isolated by paperelectrophoresis (SP-2 and 2 fragments, SP-2a/b). Sequencing results identified the last 7 residues of Tm-7. The other fraction was composed of SP-1 (31 residues). Direct sequencing resolved the identity of the first 10 residues (including Trp at position 7 of Tm-7) and of the 13th. To obtain further information on its structure, SP-1 was digested with chymotrypsin. Four fragments were formed and isolated, but limited sequence data were collected. Detailed information on SP-peptides is given in VII-I.6.4.2, table VII-I.7 and figs.VII-I.12/13.

Table VII-I.7 Amino acid composition of SP-peptides of Tm-7 and of chymotryptic (Ch-)peptides of Tm-7 SP-1.

| Amino acid | SP-1 c) | SP-2 | (SP-2a) | (SP-2b) | Total Tm-7 | SP-1 Ch-a | SP-1 Ch-b | SP-1 Ch-c | SP-1 Ch-d |
|---|---|---|---|---|---|---|---|---|---|
| Asx | 3.24(3) | 2.01(2) | 1.96(2) | | 5 | 0.94(1) | | | 2.15(2) |
| Thr d) | 3.95(4) | | | | 4 | | 0.95(1) | 2.03(2) | 1.04(1) |
| Ser d) | 2.15(2) | 1.13(1) | 0.95(1) | | 3 | 1.00(1) | 1.07(1) | | 0.33 |
| Glx | 3.11(3) | 1.11(1) | 1.09(1) | | 4 | | | | 2.73(3) |
| Pro | 1.09(1) | | | | 1 | | | | 0.89(1) |
| Gly | 1.89(2) | | | | 2 | 1.23(1) | | | 1.11(1) |
| Ala | 3.04(3) | | | | 3 | 0.23 | 2.01(2) | 1.17(1) | 0.21 |
| Val | | | | | - | | | | |
| Met e) | 0.99(1) | 0.43(1) | 0.38(1) | | 2 | | | | 0.64(1) |
| Ile f) | 2.71(3) | 0.90(1) | | 0.96(1) | 4 | 1.15(2) | | | 1.01(1) |
| Leu | 3.06(3) | | | | 3 | | | 2.00(2) | 0.96(1) |
| Tyr | 1.02(1) | | | | 1 | | 0.95(1) | | |
| Phe | 1.07(1) | | | | 1 | 0.84(1) | | | |
| His | | | | | - | | | | |
| Lys | 1.98(2) | | | | 2 | | 1.03(1) | 0.97(1) | |
| Arg | | 0.86(1) | | 1.04(1) | 1 | | | | |
| Trp | 1.70(2) | | | | 2 | n.d.(1) | | | n.d.(1) |
| No. of residues | 31 | 7 | (5) | (2) | 38 | 7 | 6 | 6 | 12 |
| Relative yield a) | 91% | 40% | 26% | 20% | | ? | 15% ( 8%) | 26% (13%) | 42% |
| $NH_2$-terminus | Gly | Ser | Ser | Ile | | | Thr | Leu | |
| Purification g) | precipitation | HVE | HVE | HVE | | precipitation Am. | HVE Am. | HVE Am. | HVE |

See also legends to tables VII-I.1 and 2.

c) Amino acid analyses of SP-1 were performed after hydrolysis for 24 and 72 h.

d) Ser and Thr values in SP-1 as found after extrapolation to zero time.

e) Low values of Met were due to partial destruction by oxidation.

f) The value given under SP-1 was found after 72 h hydrolysis; after 24 h a value of 2.37 was found. Stoichiometric value in SP-1 Ch-a is low because of incomplete hydrolysis after 24 h of the Ile-Ile bond present.

g) The SP-digest gave a precipitate on acidification; the acid-soluble fragments were purified by paperelectrophoresis (HVE) at pH 3.6. Also the Ch-digest of SP-1 contained acid-insoluble material, which was fractionated on Aminex A-5 (Am.); the soluble peptides were purified by HVE at pH 3.6.

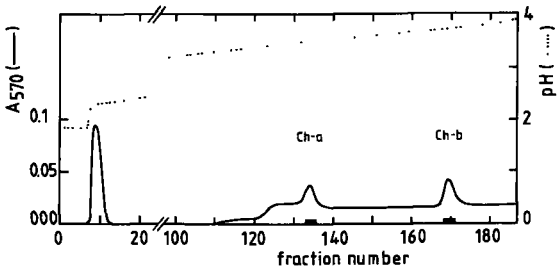



Fig.VII-I.12. Fractionation of two Ch-peptides of Tm-7 SP-1. After fractionation of the acid-soluble part of the Ch-digest by HVE, a mixture of 2 peptides resulted that was chromatographed on Aminex A-5. The column (16 x0.9 cm) was eluted at 50°C by two steps gradient elution as described for fig. VII-I.2, at a flow rate of 9.5 ml/h. Buffer A: 90 ml 0.05 M pyridinium acetate, pH 2.4; buffer B: 300 ml 0.2 M pyridinium acetate, pH 3.1; buffer C: 2.0 M pyridinium acetate, pH 5.0; flow rate: 3.9 ml/h. Fractions of 1.2 ml were collected; 34% of the effluent was consumed by on line detection of peptides in the PA (A570:——). Fractions were pooled as indicated.

VII-I.6.2.3 Isolation and sequence determination of Th-peptides

From the digest, 4 pure peptides were isolated by paperelectrophoresis, comprising only 21 of the 38 residues of Tm-7, of which a tetrapeptide was fully sequenced to provide an additional overlap between Tm-7 T-1 and T-2. Details in VII-I.6.4.3, table VII-I.6 and fig.VII-I.13.

VII-I.6.2.4 Isolation and sequence determination of CNBr-peptides

Cleavage with CNBr yielded the expected 3 fragments, which were isolated as pure products (table VII-I.8). N-terminal CNBr-1 (29 residues) was not sequenced. CNBr-2 was sequenced and shown to be identical to fully structured CN-7 (4 residues), while the 5 residues of CNBr-3 were identical to the N-terminal sequence of CN-8 (thereby providing data for correctly linking these 2 CN-peptides of atropinesterase). Details in VII-I.6.4.4 and fig.VII-I.13.

VII-I.6.2.5 Isolation and sequence determination of NBS-peptides

NBS is supposed to cleave specifically at Trp, so 3 fragments were expected. Only the C-terminal NBS-3 was isolated (16 residues) from the acid-soluble part of the digest. It was characterized by its amino

Table VII-I.8 Amino acid composition of peptides obtained after digestion of Tm-7 with trypsin, NBS and CNBr, respectively.

| Amino acid | CNBr-1 | CNBr-2 b) | CNBr-3 | Total Tm-7 | NBS-3 |
|---|---|---|---|---|---|
| Asx | 2.97(3) | | 2.08(2) | 5 | 3.86(4) |
| Thr | 3.03(3) | 1.01(1) | | 4 | 1.02(1) |
| Ser | 2.09(2) | 1.05(1) | | 3 | 0.97(1) |
| Glx | 2.01(2) | 0.94(1) | 1.08(1) | 4 | 4.09(4) |
| Pro | 0.96(1) | | | 1 | 0.96(1) |
| Gly | 1.81(2) | | | 2 | 1.12(1) |
| Ala | 3.00(3) | | | 3 | |
| Val | | | | - | |
| Met c) | 0.74(1) | 0.87(1) | | 2 | 0.74(2) |
| Ile | 2.29(3) | | 0.94(1) | 4 | 0.99(1) |
| Leu | 3.11(3) | | | 3 | |
| Tyr | 0.96(1) | | | 1 | |
| Phe | 1.05(1) | | | 1 | |
| His | | | | - | |
| Lys | 2.18(2) | | | 2 | |
| Arg | | | 0.90(1) | 1 | 0.88(1) |
| Trp | 1.30(2) | | | 2 | |
| No. of residues | 29 | 4 | 5 | 38 | 16 |
| Relative yield a) | 33% | 27% | 33% | | 35% |
| $NH_2$-terminus | n.d. | Thr | Glx | | Gly |
| Purification d) | precipitation | HVE | HVE | | HVE |

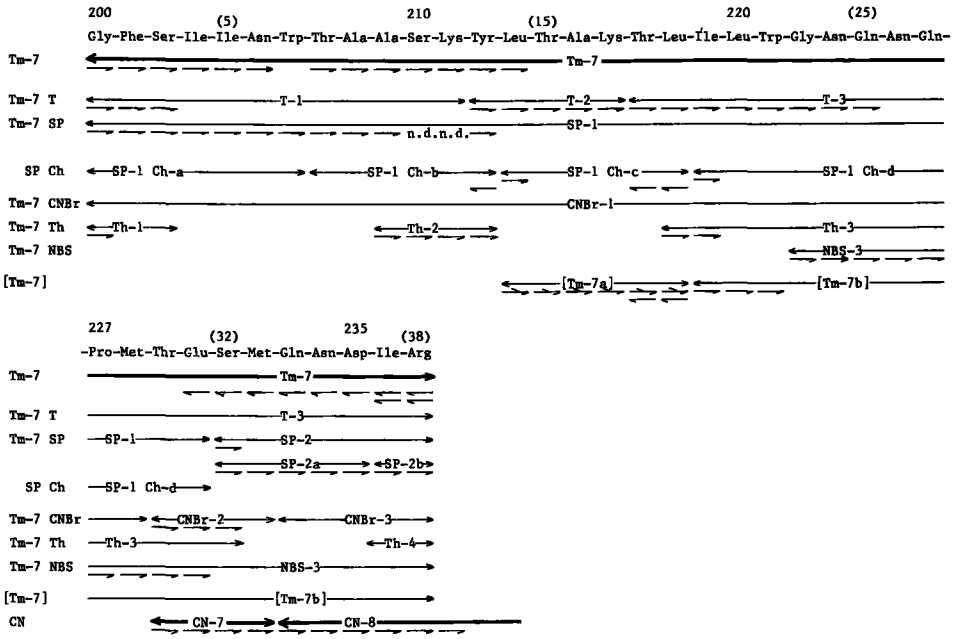
See also legends to table VII-I.1
b) CNBr-2 was detected at 2 positions after paperelectrophoresis (HVE); both eluates contained the pure peptide.
c) In CNBr-peptides Met was determined as Hse.
d) Peptides were purified by precipitation with acid and HVE at pH 3.6 of the supernatant.

acid composition (table VII-I.8) and by the sequence of the first 9 residues. This confirmed the identity of 3 positions of T-3, established the nature of 6 more and gave an overlap with CNBr-2. Details in VII-I.6.4.5 and fig.VII-I.13.

### VII-I.6.3 SEQUENCE ASSEMBLY OF Tm-7

Sequence results of Tm-7 and of fragments of Tm-7 unambiguously elucidated the complete amino acid sequence. A survey is given in fig.VII-I.13. As only 3 T-peptides were formed, which were all 3 isolated and characterized, and as T-1 was shown to be the N-terminal T-peptide because it corresponded with the first 12 residues in Tm-7, the ranking of the T-peptides was already clear. It was fully confirmed by the other data. Sequence analysis of Tm-7 combined with sequence results of Tm-7 SP-1 showed the N-terminal sequence of Tm-7 up to residue Tyr 13 in directly positioned residues. T-1 and T-2 were connected by this

Fig.VII-I.13. Amino acid sequence of Tm-7 constructed from T-, SP-, SP Ch-, CNBr-, Th-, NBS-peptides of Tm-7 and from [Tm-7a] and [Tm-7b].




Peptides are designated as described at fig.VII-I.1

sequence, T-2 extending the sequence to Lys 17. T-2 and T-3 were overlapped by SP-1 Ch-c both by composition, N-terminus and CPA results, while an aspecifically formed, fully sequenced peptide, [Tm-7a] (VII-V.2), confirmed this overlap by sequence ([Tm-7a] was identical to Ch-c as for composition, N-terminus and CPA-cleavage). Positioned residues in T-3 extended the sequence to residue 25, overlapping NBS-3 by 3 residues. NBS-3 extended the sequence to Glu 31 in directly positioned residues; this Glu was supposed to be the C-terminus of SP-1. The remaining part of NBS-3 was identical to SP-2; the sequence of this part was elucidated from the data on SP-2. The sequence of SP-2 was deduced from the completely structured fragments SP-2a and SP-2b. These fragments of SP-2 could be placed into position because SP-2b contained Arg, the C-terminal residue of Tm-7 and because SP-2a contained the same N-terminal residue as SP-2. Directly positioned residues of NBS-3 and SP-2 were overlapped by CNBr-2 of known sequence

(from direct determination and from the identical peptide CN-7 which was sequenced completely). Results obtained after hydrolysis of intact Tm-7 with CPA and CPB were consistent with the C-terminal sequence as found with CNBr-2, SP-2a and SP-2b and confirmed the presence of 1 Glu, 1 Asp and 2 amides at the C-terminal side. The sequence of the last 5 residues was confirmed by the N-terminal sequence of CN-8. All peptides of Tm-7 not mentioned confirmed the amino acid sequence by composition and partially by sequence. The assignment of the amides present in Tm-7 as determined in Tm-7, fragments of Tm-7, in CN-8 or in fragments of CN-8 is summarized in table VII-VI.

VII-I.6.4 Experimental details

VII-I.6.4.1 Sequence determination of Tm-7. Prior to the first cycle of the dansyl-Edman procedure, demaleylated Tm-7 was treated with Braunitzer's reagent (III.1.10). The degradation identified 13 of the first 14 residues. In position 4, DNS-Ile was found beside DNS-Ile-Ile (III.1.10), a spot which disappeared after prolongation of the hydrolysis time from 10 to 66 h. Asx in position 6 in Tm-7 was identified as PTH-Asn. No residue was found at position 7, but the purple colour of the solution noticed after incubation with TFA suggested the presence of Trp.
With CPA and CPB the presence of 1 Asp, 1 Glu and 2 amides was shown at the C-terminal side as given in table VI.2.

VII-I.6.4.2 Tryptic digestion of Tm-7, fractionation and sequence determination of T-peptides. Tm-7 was digested with trypsin (table III.1) and fractionated on Sephadex G-15 (fig.VII-I.9) giving T-2 (pool II), which was sequenced completely. After chromatography of pool I on Sephadex G-25 in 5% acetic acid (fig.VII-I.10) and endgroup determination of alternating fractions between 80 and 110, T-3, comprising 21 residues including Arg, was recognized in the indicated pool. The first 8 residues were sequenced by dansyl-Edman procedure. The N-terminal peptide T-1 was identified in 100 of 110 impure fractions starting at fraction number 100. Chromatography of a tryptic digest of 50 nmol of Tm-7, dissolved in 70% acetic acid, on Sephadex G-25 in 33% acetic acid, gave pure T-1 (pool III, fig.VII-I.11) according to the amino acid composition and the sequence of the first 3 residues. Pools I and II contained aggregates of tryptic peptides as found after amino acid analysis and by determination of the N-terminal sequence. The tryptic peptides could only be isolated by the methods mentioned above. Many other attempts gave no or only minor separation, which was not caused by incomplete hydrolysis of the peptide bonds at the C-terminal side of Lys as incubation of the digest with CPB liberated 2 mol Lys and 1 mol Arg per mol Tm-7. HVE and chromatography on columns of Sephadex and Bio-Gel P-4 failed despite the addition of a detergent or an organic solvent (1% SDS or 1% Triton

X-100 or 20% n-propanol) to the eluent 0.01 M Tris.HCl, pH 8.0. In all cases aggregation of T-peptides seemed to occur, probably due to the hydrophobic character of the fragments, and the allied problem of bad solubility. The amino acid compositions of the Tm-7 T-peptides (determined after 24 h of hydrolysis) are given in table VII-I.6; the low total integral number found for Ile (3) based on the stoichiometric values of T-1 (1.03) and T-3 (1.95), corresponded well with the low value (2.95) found for Ile in Tm-7 after hydrolysis for 24 h. The Trp-residue in T-1 was not determined due to the omission of 2% TGA during hydrolysis. The sequence results are summarized in fig.VII-I.13.

VII-I.6.4.3 Digestion of Tm-7 with S. protease, fractionation and sequence determination of SP-peptides. Digestion of

Tm-7 (133 nmol) with S.protease was as described in III.1.7.3. Attempts to dissolve the digest in pH 3.6 buffer gave an acid-soluble and an acid-insoluble fraction. The former was subjected to HVE at pH 3.6 (4000 V, 90 min) which yielded the C-terminal peptide SP-2, comprising residues 32 to 38, and 2 fragments, SP-2a and SP-2b, together having the same amino acid composition as SP-2. Of SP-2 only the N-terminal residue was identified, while SP-2a and SP-2b were fully sequenced. The acid-insoluble fraction, which was soluble at pH 8.0, contained the pure N-terminal peptide SP-1 (31 residues). Sequencing of the first 10 residues and identification of Tyr at position 13 confirmed the N-terminal sequence of Tm-7 and definitely showed Trp to occupy position 7 by identification of the DNS-derivative. The production of SP-1, SP-2a and SP-2b, which had taken place by the specific action of S.protease, showed the acidic character of Glu in position 31 and Asp in position 36. The presence of Asp in position 36 and the net charge found for SP-2a, being -1 according to Offord, assigned Glx and Asx in positions 34 and 35 in Tm-7 both as amides, which is consistent with the net charge of NBS-3 being -1 (VII-I.6.4.6) and with the sequence results of CN-8 (VII-II.3).
To provide an overlap from T-2 to T-3, Tm-7 SP-1 was digested with chymotrypsin (III.1.7.5) at pH 8.0, giving a soluble and an insoluble fraction at pH 3.6. The former, fractionated by HVE at pH 3.6 (4000 V, 2h), gave pure SP-1 Ch-d, which comprised residues 20 to 31 as judged by the amino acid composition, and a mixture of Ch-b and Ch-c in a ratio of 1.0:1.7 as judged by amino acid analysis and from the amounts of Tyr and Thr + Leu, released by CPA. These peptides, eluted from paper, were purified by fractionation on Aminex A-5 (fig.VII-I.12). The amino acid compositions were determined. In Ch-c (residues 14 to 19) Leu was shown to be N-terminal, while digestion with CPA released Thr and Leu in equimolar amounts (table VII-I.10). Ch-c represented the desired fragment overlapping T-2 and T-3. According to the amino acid analysis this peptide was identical to a fully structured fragment, [Tm-7a] (VII-V.2). The acid-insoluble fraction delivered, after redigestion with chymotrypsin and fractionation on Aminex A-5, pure SP-1 Ch-a according to the amino acid composition. The peptide was eluted at pH 2.9 (as described in fig.VII-I.12; result not shown). The 4 Ch-peptides comprised all constituting amino acids of Tm-7 SP-1. The

amino acid compositions and other details of SP-peptides and SP Ch-peptides are given in table VII-I.7.10. All sequence results are included in fig.VII-I.13.

VII-I.6.4.4 Thermolytic digestion of Tm-7, fractionation and sequence determination of Th-peptides. The peptide mixture obtained after digestion of Tm-7 (150 nmol) with thermolysin (III.1.7.4) was fractionated by HVE at pH 3.6 (4000 V, 80 min) giving 4 pure peptides with amino acid compositions as shown in table VII-I.6. Partially or fully structured peptides Th-1, Th-2, Th-3 and Th-4 (fig.VII-I.13) confirmed partial sequences of Tm-7.

VII-I.6.4.5 CNBr cleavage of Tm-7, fractionation and sequence determination of CNBr-peptides. CNBr cleavage of Tm-7 resulted in the isolation of an acid-soluble and an insoluble fraction. The latter contained pure CNBr-1, comprising residues 1 to 29 as judged by amino acid analysis. Separation of the acid-soluble peptides by HVE at pH 3.6 (4000 V, 120 min) gave pure CNBr-2 and CNBr-3 with well-fitting amino acid compositions. Of CNBr-2, present in 2 bands on the electropherogram, probably due to the occurrence of Hse next to Hse-lactone, 3 residues were sequenced. It comprised residues 30 to 33 of Tm-7 and was identical to fully structured CN-7.
CNBr-3 confirmed the C-terminal sequence of Tm-7. The amino acid compositions of CNBr-peptides are given in table VII-I.8. Sequence results of CNBr-2 are shown in fig.VII-I.13.

VII-I.6.4.6 Cleavage of Tm-7 with NBS, fractionation of NBS-peptides and sequence determination of the C-terminal fragment. After cleavage of Tm-7 (135 nmol) with NBS again an acid-soluble and an insoluble fraction were obtained (III.1.7.6). After fractionation of the acid-soluble part by HVE at pH 3.6 (4000 V, 2½ h) only the Arg-containing peptide NBS-3 (residues 23 to 38) was detected after staining with fluorescamine and phenanthrenequinone. The amino acid composition was established (table VII-I.8) and the sequence of the first 9 residues (fig.VII-I.13). According to Offord NBS-3 had a net charge of -1, indicating the presence of 1 basic and 2 acidic residues and the presence of 6 amides; 4 of these amides were found at positions 24 to 27 in Tm-7, by identification of the PTH-derivatives. The character of Asx and Glx at the C-terminal side of NBS-3 was deduced from results of SP-peptides (VII-I.6.4.3). The acid-soluble fraction in which the N-terminal part of Tm-7 was tentatively identified, was not investigated further.

## VII-I.7 Tm-8 (35 RESIDUES)

### VII-I.7.1 SEQUENCE DETERMINATION OF Tm-8

With the dansyl-Edman degradation the identity of 11 residues was determined. CPA and CPB released the same amino acids from Tm-8, and

in comparable relative amounts, as from intact atropinesterase (compare tables IV.2 and VI.2), indicating that this peptide was the C-terminal fragment as was already suggested by the absence of Arg.

VII-I.7.2 SECONDARY CLEAVAGES OF Tm-8

Digestions were performed with trypsin, pepsin and CNBr. All residues present in Tm-8 were identified in 4 tryptic (T-) and 5 peptic (P-) peptides. Structural analysis of these peptides confirmed the N-terminal sequence of Tm-8 and established the C-terminal sequence of 15 residues. The CNBr-peptides formed showed that Tm-8 comprised CN-9 and CN-10. For completing the sequence determination of Tm-8, the structure of a peptic peptide of atropinesterase, identical to Tm-8 P-2, was used.

VII-I.7.2.1 Isolation and sequence determination of T-peptides

The digests of Tm-8 were fractionated by gelfiltration followed by paperelectrophoresis, eventually resulting in the isolation of all 4 T-peptides. T-1 with 5, and T-2 with 2 residues were both obtained in good yield; they were sequenced completely. T-3 (21 residues) was obtained in reasonable yield. The first 3 residues were sequenced. The sequence results of T-1, T-2 and T-3 confirmed the N-terminal sequence found with Tm-8. CPA/CPB identified the last 4 residues of T-3. T-4 (7 residues), isolated in high yield, was only partly sequenced. Details are given in VII-I.7.4.1, table VII-I.9.10 and figs.VII-I.14/16.

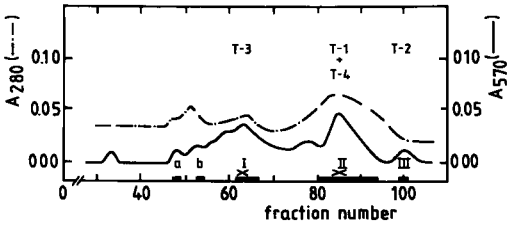



Fig.VII-I.14. Fractionation of T-peptides of Tm-8. A column of Sephadex G-50 sf was eluted with 50% acetic acid at 18 ml/h. Fractions of 2.5 ml were collected; 17% of the effluent was consumed by on line detection of peptides in the PA (A570:——). A280 (-·-·-) of the fractions was measured. Pooling was as indicated. Crosses refer to fractions used for sequence elucidation.

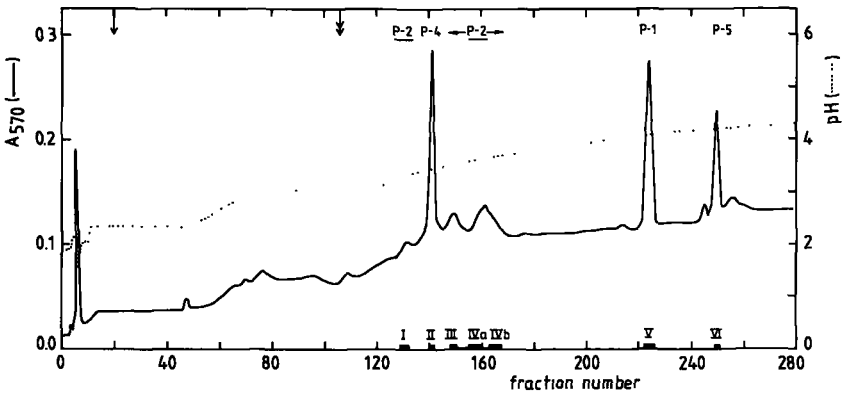



Fig.VII-I.15. Fractionation of P-peptides of Tm-8. A column of Aminex A-5 (19 x 0.9 cm) was eluted at 50°C with 40 ml 0.05 M pyridinium acetate, pH 2.4 (buffer A), followed by two steps gradient elution as described for fig. VII-I.2, at a flow rate of 11.4 ml/h. Buffer A: 90 ml; buffer B: 350 ml 0.2 M pyridinium acetate, pH 3.1; buffer C: 2.0 M pyridinium acetate, pH 5.0; pumping rate: 4.9 ml/h. ↓, ↓ :start 1st and 2nd gradient. Fractions of 1.3 ml were collected; 28% of the effluent was consumed by on line detection of peptides in the PA (A570:——). Fractions were pooled as indicated; the underlined peptide was obtained together with contaminating material.

Table VII-I.9 Amino acid composition of T- and P-peptides of Tm-8.

| Amino acid | T-1 | T-2 | T-3[b] | T-4 | P-1 | P-2[b] | P-3 | (P-3a) | P-4 | P-5 | Total Tm-8 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | | | 2.25(2) | 1.98(2) | | 1.28(1) | 1.09(1) | 1.07(1) | 1.99(2) | | 4 |
| Thr | | | 0.30 | | | | | | | | |
| Ser | | | 1.01(1) | | | 1.05(1) | 0.26 | | | | 1 |
| Glx | | | 2.96(3) | 1.14(1) | | 1.28(1) | 2.02(2) | 0.93(1) | 1.20(1) | | 4 |
| Pro | 1.03(1) | | 1.09(1) | 0.21 | 0.99(1) | | 1.09(1) | | | | 2 |
| Gly | | | 1.99(2) | | | 2.27(2) | 0.40 | | 0.33 | | 2 |
| Ala | 1.98(2) | 0.98(1) | 1.23(1) | 0.30 | 2.98(3) | 0.56 | | | 0.97(1) | | 4 |
| Val | | | 0.95(1) | | | 0.33 | | | 0.91(1) | | 1 |
| Met | | | 1.95(2) | | | n.f.(1) | 0.81(1) | | | | 2 |
| Ile | | | 1.05(1) | | | 0.82(1) | | | | | 1 |
| Leu | 0.96(1) | | 0.42 | 1.87(2) | 1.06(1) | 0.37 | | | 1.00(1) | 1.01(1) | 3 |
| Tyr | | | 0.87(1) | | | 0.72(1) | | | | | 1 |
| Phe | | | 2.45(3)[c] | 0.99(1) | 1.04(1) | 1.65(2) | | | 0.24 | 0.98(1) | 4 |
| His | | | 1.02(1) | | | 0.92(1) | | | 0.21 | | 1 |
| Lys | 1.03(1) | 1.02(1) | 1.14(1) | 1.03(1) | 1.94(2) | | | | 0.93(1) | 1.01(1) | 4 |
| Arg | | | | | | | | | | | |
| Trp | | | 0.85(1) | | | | 0.16(0-1) | 0.50(1) | | | 1 |
| No. of residues | 5 | 2 | 21 | 7 | 8 | 11 | 5-6 | (3) | 7 | 3 | 35 |
| Relative yield a) | 35% | 29% | | | 50% | 7% | | | 50% | 57% | |
| | 76% (15%) | 22% | 18% | 93% (34%) | | | 25% | 11% | 44% | | |
| $NH_2$-terminus | Ala | Ala | Phe | Asx | Ala | Ile | Trp | | Val | Phe | |
| purification d) | | | | | | | | | | | |
| A | G-50+HVE | G-50+HVE | | | Am. | Am. | HVE | HVE | Am. | Am. | |
| B | G-50+HVE | G-50 | G-50 | G-50+HVE | | | | | HVE | | |
| Pool (B) | II | III | I | II | V | I+III+IV | | | II | VI | |

See also legends to tables VII-I.1 and 2 (upper relative yields refer to purification system A, lower to B).

c) A low value for Phe was found, as was the case in Tm-8 as well.

d) T-peptides were purified by gelfiltration on Sephadex G-50 sf (G-50), either with 5% (A) or 50% (B) acetic acid and subsequent paperelectrophoresis at pH 3.6 (HVE). P-peptides were purified by ion-exchange chromatography on Aminex A-5 (Am.) or HVE or both, as indicated. Pool nrs. refer to figs.VII-I.14/15.

VII-I.7.2.2 Isolation and sequence determination of P-peptides

Peptic digests were fractionated by ion-exchange chromatography (fig. VII-I.15) or by paperelectrophoresis. Eventually, all 5 P-peptides were isolated, albeit P-2 (11 residues) not very pure and in low yield. The octapeptide P-1 was identified as the N-terminal fragment because of sequence and composition similarity with the N-terminus of Tm-8, and with T-1 and T-2. C-terminal tripeptide P-5, the heptapeptide P-4 and the hexapeptide P-3 were fully sequenced. Of P-2 only the first and the last 2 residues were determined. Further details are given in VII-I.7.4.3., tables VII-I.9-10 and fig.VII-I.16.

VII-I.7.2.3 Isolation and sequence determination of CNBr-peptides

The 3 fragments could not be separated. Sequence information was obtained by dansyl-Edman degradation of the mixture. The results were consistent with the N-terminal sequences of Tm-8, CN-9 and CN-10.

Table VII-I.10 Amino acids released by leucinaminopeptidase (LAP) or carboxypeptidase A and/or B (CPA/B) from T-peptides, P-peptides, one SP- and one Ch-peptide of Tm-fragments (Tm-2, Tm-7 and Tm-8) and from one P-peptide of atropinesterase (P-B).

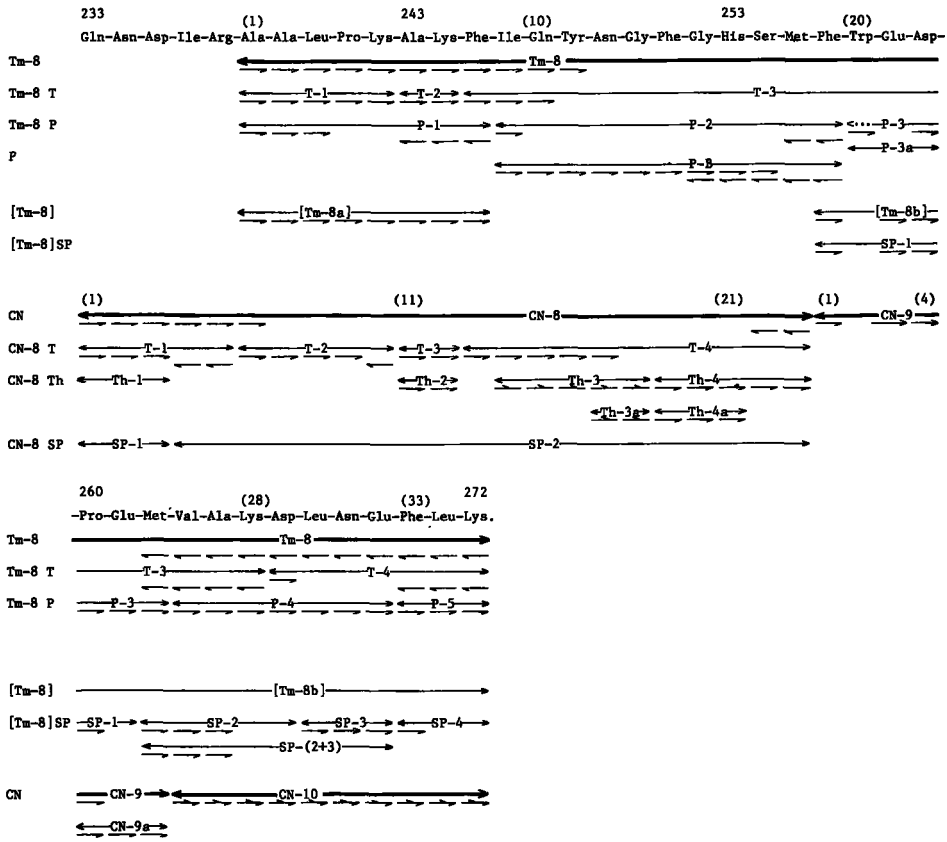
| Peptide | Enzyme | Incubation time(h) | Amino acids released[a] |
|---|---|---|---|
| Tm-2 P-2 | | | ...-Gly-Tyr-Thr-Asp-Thr-Ser-Arg |
| | CPA/B | 18 | 0.2 0.7 0.8 |
| Tm-2 SP-1 | | | Tyr-Val-Glu |
| | LAP | 2 | 1.0 1.0 1.0 |
| Tm-7 SP-1 Ch-b | | | Leu-Thr-Ala-Lys-Thr-Leu |
| | CPA | ¼ | 1.0 1.0 |
| Tm-8 T-3 | | | ...-Pro-Glu-Met-Val-Ala-Lys |
| | CPA/B | 2 | 0.6 0.7 0.8 0.9 |
| Tm-8 T-4 | | | Asp-Leu-Asn-Glu-Phe-Leu-Lys |
| | CPA | ½ | 0.8 0.9 1.0 |
| Tm-8 P-1 | | | Ala-Ala-Leu-Pro-Lys-Ala-Lys-Phe |
| | CPA/B | ½ | 1.0 1.1 0.9 |
| Tm-8 P-2 | | | Ile-Gln-Tyr-Asn-Gly-Phe-Gly-His-Ser-Met[b]-Phe |
| | CPA | ½ | 0.5 1.0 |
| P-B | CPA/B | 4 | 0.3 0.8 1.0 |
| | | 18 | 0.7 0.5 0.6 0.9 1.2 |
| Tm-8 P-4 | | | Val-Ala-Lys-Asp-Leu-Asn[c]-Glu |
| | LAP | 18 | 1.0 1.0 1.0 1.1 1.0 0.8 1.1 |

Residues are ranked according to their sequence in the peptides.
a) Residues released are reported as mol of residue per mol of peptide.
b) Met was determined as Met sulfoxide.
c) Asn co-eluted with Ser and was calculated by using the standard value of Ser. A low ratio A570/A440 was found indicating the presence of an amide.

## VII-I.7.3 SEQUENCE ASSEMBLY OF Tm-8

Sequence results of Tm-8 and fragments of Tm-8, combined with additional evidence from CN-8, CN-9 and CN-10 (VII-II.3.4.5) and from P-B (VII-IV) established the whole sequence of Tm-8 as is shown in the survey in fig.VII-I.16.

Fig.VII-I.16. Amino acid sequence of Tm-8 constructed from Tm-8, Tm-8b, CN-8, CN-9, CN-10 and fragments of these peptides and from one P-peptide of atropinesterase, P-B.




Peptides are designated as described at fig.VII-I.1

· · · In P-3 the same phenomenon was observed as given for T-6 P-3 and P-6, see fig.VII-I.8

VII-I.7.4 Experimental details

VII-I.7.4.1 Sequence determination of Tm-8. The first 11 residues were sequenced. Starting at Pro 4, the newly formed DNS-amino acid was slightly contaminated with the DNS-amino acid of the previous step. Despite this "ragged" degradation the amino acid sequence up to the 11th position could be determined unambiguously. With CPA and CPB the last 11 residues were released, showing 1 Asp, 1 Glu and 1 Asn to be present (table VI.2).

VII-I.7.4.2 Tryptic digestion of Tm-8, fractionation and sequence determination of T-peptides. Tm-8 was digested with trypsin as described in chapter III. In a first experiment (225 nmol) no separation was obtained on a column (52 x 2cm) of Sephadex G-50 sf in 5% acetic acid. When material of all peptide-containing fractions was fractionated by HVE at pH 3.6 (4000 V, 105 min) 2 pure peptides T-1 and T-2 were obtained, which were sequenced completely. T-3 and T-4 were only tentatively identified in low yield in impure eluates. In a second experiment (175 nmol) a column (104 x 2.3 cm) of Sephadex G-50 sf and 50% acetic acid were used (fig.VII-I.14). Amino acid analysis of pool III and of some fractions of pool I (indicated by a cross) showed T-2 and T-3 to be present in homogeneous form. Because of the low amount of T-3, only the sequence of the first 3 residues was determined; with CPA and CPB the last 4 residues were identified. In the central fractions of pool II, a mixture of T-1 and T-4 could be characterized. HVE of pool II at pH 3.6 (4000 V, 90 min) gave pure T-1 and T-4. Of T-4 only the N-terminal residue Asp was determined, while with CPA the last 3 residues were released. Pools a and b contained aggregated T-peptides.

VII-I.7.4.3 Peptic digestion of Tm-8, fractionation and sequence determination of P-peptides. In a first experiment the digest of 135 nmol Tm-8 was fractionated on a column of Aminex A-5 at 50 °C (fig. VII-I.15). P-1, P-4 and P-5 were isolated in pure state. P-2 was tentatively identified in many fractions at different pH values. Pool I contained P-2 in fairly homogeneous form but according to the amino acid analysis some contaminating amino acids were present. Amounts were too low for additional purification steps. Of pure P-1 the first 3 residues were sequenced, subsequent residues were not found probably due to extraction of the remaining peptide by ethyl acetate. From the C-terminus 3 residues were released by CPA/CPB which agreed with the residues found for the positions 6, 7 and 8 in Tm-8. Of P-2 the first residue Ile was detected with dansylchloride, while Met and Phe were released by CPA. Of P-4 and P-5 the whole sequence was established. Digestion of P-4 with LAP showed the presence of 1 Asp, 1 Asn and 1 Glu, which was consistent with the CPA and CPB results of Tm-8 (table VI.2). In a second experiment the peptic digest of 60 nmol Tm-8 was fractionated by HVE at pH 3.6 (4000 V, 120 min). Detection was on guidestrips with fluorescamine and His reagent. No His-containing peptides were detected (1 His is present in P-2). One eluate contained a mixture of P-1 and P-5 according to the amino acid analysis and endgroup determination with dansylchloride; yields were

14 and 35%, respectively. P-3 was found with a well-fitting amino acid composition but for a low ratio number of Trp, while some contaminating Gly was found. During sequencing of P-3 the same problem was encountered as with Tm-6 P-3 and P-6 (VII-I.5.4.2). Despite the low ratio number of Trp in Tm-8 P-3, Trp was clearly identified as N-terminal residue, while after each Edman degradation cycle 2 DNS-derivatives were detected, one being in accordance with Trp as N-terminal residue, the other corresponding with the residue one Edman cycle ahead. The sequence of the residues 3, 4, 5 and 6 of P-3 (assuming Trp to be N-terminal), was established with certainty. P-3a, representing the first 3 residues of P-3, was identified in low yield (11%). P-1 and P-5 were recognized in a peptide mixture. P-2 was not found in any of the eluates (His-reagent; amino acid analysis). Pure P-4 was isolated in good yield.
Amino acid compositions, relative yields and N-terminal residues of all P-peptides of Tm-8 are summarized in table VII-I.9; results obtained with CPA and CPB are given in table VII-I.10.

VII-I.7.4.4 CNBr-cleavage of Tm-8 and sequence determination of the CNBr-peptides. Tm-8 was digested with CNBr as described in III.1.7.7. The amino acid sequence of 3 CNBr-peptides was determined simultaneously because the available amount (10 nmol) was too small for further fractionation steps. Beside the N-terminus Ala-Ala of Tm-8, the sequence Phe-Trp was found corresponding with the N-terminus of CN-9, and Val-Ala which is identical to the N-terminus of CN-10. No free amino acids were present as was shown by amino acid analysis without acidic hydrolysis. These results suggested the presence of the sequences Met-Phe-Trp and Met-Val-Ala in Tm-8.

Table VII-I.11 Amino acids released by leucinaminopeptidase (LAP) or carboxypeptidase A and/or B (CPA/B) from T- and P-peptides of Tm-6 and/or Cm-6a.

| Peptide | Enzyme | Incubation time | Amino acids released[a] |
|---|---|---|---|
| Tm-6 T-1[d] | | | Leu-Val-Leu-Ile-Ser-Thr-Ala-Leu-Lys |
| | CPA/B | 10 min | 0.2 0.2 0.9 1.1 1.0 |
| | CPA/B | 6 h | 0.9 0.9 1.0 1.1 1.1 2.6 1.0 |
| | | | f) f) |
| Tm-6 T-2[e] | | | ...-Pro-Val-Leu-Glu-Trp-Val-Tyr-Asp-Thr-Val-Leu-Gln-Lys |
| | CPA/B | 45 min | 0.5 0.5 0.4 0.4 0.8 1.1 1.3 1.3 |
| | CPA/B | 21 h | 1.0 1.0 0.9 0.9 1.6 2.3 2.0 1.3 |
| Tm-6 P-3 | | | Trp-Val-Tyr-Asp-Thr-Val-Leu |
| | CPA | 18 h | 0.9 0.9 1.0 1.0 1.0 |
| Tm-6 P-5 | | | Phe-Ala-Lys-Glu |
| | CPA | 18 h | 0.3 0.4 1.0 1.0 |
| Tm-6 T-4[g] | | | Glu-Trp-Val-Ala-Ala-Pro-Gly-Lys |
| | LAP | 18 h | 0.9 1.1 1.0 1.1 |
| | CPA/B | 18 h | 0.3 0.7 |
| Tm-6 T-5 | | | *His-Asp-Asn-Gly-Met-Ala-Lys* |
| | CPA/B | 18 h | 0.8 0.9 1.0 |
| | | | e) |
| Cm-6a T-6 | | | Asn-Leu-Lys |
| | LAP | 18 h | 0.6 1.0 1.0 |
| Tm-6 T-8 | | | His-Val-Trp-Leu-Ser-Ala-Ala-Arg |
| | CPA/B | 12 min | 1.0 0.9 0.9 1.9 1.0 |
| Tm-6 P-7 | | | ...-Pro-Lys-His-Val-Trp-Leu-Ser-Ala-Ala-Arg |
| | CPA/B | 2½h | 1.0 0.9 1.1 1.0 1.0 2.1 0.9 |

See also legends to table VII-I.10.
d) and e) Incubations were performed with Cm-6a T-1 and/or Tm-6 T-1. See VII-I.5.4.1d.
e) T-2 was contaminated with T-3 (C-terminal sequence: -Ser-Glu-Phe-Ala-Lys); the following amino acids not derived from T-2 were found: after 21 h: Ser(0.6) Ala(0.5) Phe(0.3).
f) Asn and Gln co-eluted with Thr, but were recognized on their high ratio A440/A570. Values were calculated with the standard value of Thr (ca.1.4x the value of Asn and Gln)
g) LAP-digestion was performed with T-4 from Tm-6 and Cm-6a (combined results are given). Only Cm-6a T-4 was incubated with CPA/B.

## VII-II SEQUENCE DETERMINATION OF FIVE CN-PEPTIDES

### VII-II.1 CN-1 (26 RESIDUES)

#### VII-II.1.1 SEQUENCE DETERMINATION OF CN-1

Direct sequence determination with the dansyl-Edman procedure revealed the identity of the first 10 residues. These were identical to the corresponding amino acids in Tm-1 and in the intact protein. This identified CN-1 as the N-terminal CN-fragment of atropinesterase. The composition of the remaining part of CN-1 and the nature of the 4 residues released from CN-1 by CPA and CPB were in agreement with the next 16 residues in intact atropinesterase (see IV.6).

#### VII-II.1.2 SECONDARY CLEAVAGES OF CN-1

Cleavages were performed with trypsin, S.protease and thermolysin. The distribution of the 3 Lys-residues (1 penultimate, 2 next to each other; no Arg) made the tryptic digestion less informative than expected. As a consequence, digestions with 2 other enzymes were needed to establish the complete sequence.

##### VII-II.1.2.1 Isolation and characterization of T-peptides of CN-1

Fractionation of the tryptic digest by paperelectrophoresis yielded free Lys and Hse (T-2 and T-4), the 2 other T-peptides T-1 and T-3 (15 and 9 residues, resp.) and T-1b, a C-terminal tetrapeptide of T-1. T-1 was identified as the N-terminal fragment on the basis of 9 sequenced residues. T-1b was fully sequenced, and of T-3 7 residues. With these peptides, all residues in CN-1 were recovered. Details are given in VII-II.1.4.1, tables VII-II.1 and 3 and fig.VII-I.1.

##### VII-II.1.2.2 Isolation and characterization of SP-peptides of CN-1

Paperelectrophoresis of the digest resulted in 4 SP-peptides, comprising all residues of CN-1. Only SP-4, the C-terminal tetrapeptide, was sequenced to provide the missing sequence information of T-3 and the overlap with the C-terminal Hse in CN-1. On the basis of the specificity of S.protease, which cleaves after acidic residues, a number of

Table VII-II.1 Amino acid composition of T-peptides, Th-peptides and SP-peptides of CN-1.

| Amino acid | T-1 | T-2 | T-3 | (T-1b) | Th-1 | Th-2 | Th-3 | Th-4 | SP-1 | SP-2 | SP-3 | SP-4 | Total CN-1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | 2.22(2) | | 2.13(2) | 0.87(1) | 1.30(1) | 0.88(1) | | | | 0.89(1) | 3.09(3) | | 4 |
| Thr | | | | | | | | | | | | | |
| Ser | 1.03(1) | | 0.97(1) | 0.85(1) | | | 1.93(2) | | | 0.26 | 1.98(2) | | 2 |
| Glx | 2.07(2) | | 0.97(1) | | 2.04(2) | | | | 1.00(1) | | 2.05(2) | | 3 |
| Pro | 1.95(2) | | | | 2.27(2) | | | | | 1.61(2) | 0.24 | | 2 |
| Gly | | | | | | | | | | 0.30 | | | |
| Ala | 3.07(3) | | 1.07(1) | 1.25(1) | 1.14(1) | 1.12(1) | 1.13(1) | | | | 3.30(3) | 1.14(1) | 4 |
| Val | 0.75(1) | | | | 0.88(1) | | | | | 1.12(1) | | | 1 |
| Met b) | | | | | | | | | | | | 0.99(1) | 1 |
| Ile c) | 1.52(2) | | 2.23(3) | | 1.36(2) | | | 0.94(1) | | 0.83(2) | 1.76(2) | 0.98(1) | 5 |
| Leu | | | | | | | | | | | | | |
| Tyr | | | | | | | | | | | | | |
| Phe | | | | | | | | | | | | | |
| His | | | | | | | | | | | | | |
| Lys | 1.11(1) | 1.00(1) | 0.87(1) | 1.03(1) | | | 1.95(2) | 1.06(1) | | | 1.84(2) | 0.90(1) | 3 |
| Arg | | | | | | | | | | | | | |
| Trp | n.d.(1) | | | | | 0.44(1) | | | | | n.d.(1) | | 1 |
| No.of residues | 15 | 1 | 9 | (4) | 9 | 3 | 5 | 2 | 1 | 6 | 15 | 4 | 26 |
| Relative yield a) | 60% | 75% | 26% | 10% | 75% | 75% | 50% | 60% | 29% | 16% | 5% | 39% | |
| $NH_2$-terminus | Glx | Lys | Ser | Asn | Glx | Ala | Ala | Ile | Glx | n.d. | n.d. | Ala | |

See also legends to tables VII.I.1 and 2.
b) Met was present as Hse.
c) Low stoichiometric value due to the presence of Ile-Ile.

acid/amide assignments appeared possible. Details in VII-II.1.4.2 and table VII-II.1.

VII-II.1.2.3 Isolation and characterization of Th-peptides of CN-1

Fractionation by paperelectrophoresis yielded 4 sufficiently pure fragments, comprising 19 of the 26 residues of CN-1. Th-1 was recognized as the N-terminal nonapeptide because of its composition and terminal residues. Th-2, Th-3 and Th-4 (with 3, 5 and 2 residues) were fully sequenced. Th-3 provided the overlap T-1 - T-2 - T-3. Details in VII-II.1.4.3, tables VII-II.1 and 3 and fig.VII-I.1.

VII-II.1.3 SEQUENCE ASSEMBLY OF CN-1

A survey of this assmbly is given in fig.VII-I.1. The determined N-terminal sequence of CN-1 comprised the 9 sequenced residues of T-1, total Th-1 by composition and the first established residue of Th-2. Th-2, fully sequenced, and Th-3 were connected by the sequence of T-1b, which elucidated the sequence of CN-1 up to residue 17. Th-3 overlapped T-2 (Lys) and T-3 by the position of one residue which aligned T-1, T-2 and T-3 and extended the sequence to the 7th residue

of T-3 in position 23. The sequence of the last 3 residues of CN-1 was elucidated by structural analysis of SP-4. Amino acid compositions and partially or fully determined sequences of T-1, T-2, Th-1, Th-4, SP-1, SP-2 and SP-3 were consistent with the established structure of CN-1. CPA and CPB hydrolysis of CN-1 T-3 and CN-1 confirmed the sequence determined. Data on amide assignment are given in VII-II.1.4.3. Methods used to assign amides in CN-1 are summarized in table VII-VI. The sequence determined for CN-1 was in complete agreement with the results obtained with Tm-1 and with the N-terminal sequence established for the intact protein.

VII-II.1.4 Experimental details

VII-II.1.4.1 Digestion of CN-1 with trypsin, fractionation and sequence determination of T-peptides. Peptide mixtures obtained after digestion of CN-1 with trypsin (table III.1) were fractionated by HVE at pH 3.6 (2000 V, 70 min). T-1, T-1b, T-2 and T-3 were isolated in pure state as given by the amino acid compositions (table VII-II.1). One eluate contained a mixture of impure T-1a, comprising residues 1-11, and T-4, the C-terminal amino acid Hse (identified by direct amino acid analysis). Nine residues of T-1 were sequenced; the amino acids in positions 12-15 were identified by sequence analysis of T-1b; T-2, Lys, was identified as DNS-derivative. The position of 7 residues of T-3 was determined and 2 C-terminal residues were released by CPA and CPB (table VII-II.3). T-1b was produced by partial cleavage at the C-terminal side of Trp, probably due to chymotryptic-like action of trypsin.

VII-II.1.4.2 Digestion of CN-1 with S.protease, fractionation and sequence determination of SP-peptides. The digest made according to III.1.7.3 and table III.1 was fractionated by HVE at pH 6.5 (3000 V, 40 min). SP-1, SP-2, SP-3 and SP-4, comprising 1, 6, 15 and 4 residues, respectively, were isolated in a homogeneous form as judged by the amino acid analyses (table VII-II.1). Only SP-4 was sequenced (fig.VII-I.1). Because of the compositions of SP-2, -3 and -4, SP-1 should represent the N-terminal Glx of CN-1. On HVE it comigrated with Glu. This and the specificity of S.protease identified this residue as the acid, in agreement with sequencer data of atropinesterase. SP-2, representing residues 2-7, identified Asx (7) as Asp, again conform sequencer results. According to its amino acid composition SP-3 corresponded to residues 8-22 of CN-1. The hydrolysis by S.protease after residue 22 identified the Asx found for this position in T-3 as Asp. According to Offord, SP-3 had a net charge of +1. As it contained 2 basic residues, the presence of one acidic residue was indicated. Since one such residue was identified as the C-terminal Asp (22), it follows that Glx (8), Asx (12), Glx (19) and Asx (21) should all be present as the amides.

VII-II.1.4.3 Digestion of CN-1 with thermolysin, fractionation and sequence determination of Th-peptides. Digest was prepared and fractionated as mentioned for tryptic peptides. HVE at pH 3.6 gave pure Th-1, Th-2, Th-3 and Th-4 with amino acid compositions as given in table VII-II.1. Th-2, Th-3 and Th-4 were fully sequenced. Th-1 was characterized by dansylation and CPA hydrolysis giving Glx and Ala as N- and C-terminal residue, respectively. Incubation of Th-2 with LAP (table VII-II.3) showed the presence of Asn at position 12.

## VII-II.2 CN-7 (4 RESIDUES)

Of this tetrapeptide the whole sequence was established as given in fig.VII-I.13.

## VII-II.3 CN-8 (23 RESIDUES)

### VII-II.3.1 SEQUENCE DETERMINATION OF CN-8

The position of the first 6 residues of CN-8, the only Arg-containing CN-peptide of atropinesterase isolated in pure state, was determined after deblocking of the N-terminus (VI-II.2.1). Arg was found at position 5. CPA and CPB results (table VI.4) showed Ser and Hse to be the C-terminal residues.

### VII-II.3.2 SECONDARY CLEAVAGES OF CN-8

CN-8 was digested with trypsin, S.protease and thermolysin. Trypsin gave 4 fragments in which all residues of CN-8 were present. T-1, -2 and -4 were partially sequenced, the dipeptide T-3 completely. In the Th-peptides isolated, only 15 residues of CN-8 were represented. S.protease cleaved only once, after the third residue.

#### VII-II.3.2.1 Isolation and sequence determination of T-peptides

The tryptic digest of CN-8 was fractionated by paperelectrophoresis, yielding 4 pure peptides in good yields. Of T-1 the first 3 residues were identified. The remaining sequence was established by CPA/CPB studies as Ile-Arg. T-1 data agreed with the N-terminal sequence of CN-8. T-2 and T-3 were identified as internal fragments (C-terminal Lys), which appeared identical to the tryptic peptides T-1 and T-2 of Tm-8. In T-4, the identity of 6 of the 11 residues remained undetermined. Details are given in VII-II.3.4.1, tables VII-II.2/3 and fig. VII-I.16.

Table VII-II.2 Amino acid composition of T-peptides, Th-peptides and SP-peptides of CN-8.

| Amino acid | T-1 | T-2 | T-3 | T-4 | Th-1 | Th-2 | Th-3 | Th-4 | (Th-3a) | (Th-4a) | SP-1 | SP-2 | Total CN-8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | 2.06(2) | | | 1.24(1) | 1.91(2) | | 1.03(1) | | 0.91(1) | | 1.90(2) | 1.03(1) | 3 |
| Thr | | | | | | | | | | | | | |
| Ser | | | | 0.93(1) | | | | 1.03(1) | | | | 1.11(1) | 1 |
| Glx | 1.12(1) | | | 1.16(1) | 1.09(1) | | 0.98(1) | | | | 1.10(1) | 1.15(1) | 2 |
| Pro | | 1.00(1) | | | | | | | | | | 0.95(1) | 1 |
| Gly | | | | 2.04(2) | | | 1.04(1) | 1.18(1) | 1.09(1) | 1.10(1) | | 2.32(2) | 2 |
| Ala | | 2.01(2) | 1.06(1) | | | 1.02(1) | | | | | | 2.96(3) | 3 |
| Val | | | | | | | | | | | | | |
| Met b) | | | | 0.91(1) | | | | 0.84(1) | | | | n.d.(1) | 1 |
| Ile | 0.95(1) | | | 1.00(1) | | | 0.97(1) | | | | | 1.73(2) | 2 |
| Leu | | 0.97(1) | | | | | | | | | | 1.05(1) | 1 |
| Tyr | | | | 0.81(1) | | | 0.96(1) | | | | | 0.91(1) | 1 |
| Phe | | | | 1.87(2) | | | | 1.03(1) | | 1.06(1) | | 2.06(2) | 2 |
| His | | | | 1.04(1) | | | | 0.92(1) | | 0.85(1) | | 0.99(1) | 1 |
| Lys | | 1.02(1) | 0.94(1) | | | 0.98(1) | | | | | | 1.87(2) | 2 |
| Arg | 0.87(1) | | | | | | | | | | | 0.86(1) | 1 |
| Trp | | | | | | | | | | | | | |
| No. of residues | 5 | 5 | 2 | 11 | 3 | 2 | 5 | 5 | (2) | (3) | 3 | 20 | 23 |
| Relative yield a) | 58% | 68% | 66% | 37% | 23% | 14% | 54% | 34% | 29% | 34% | 12% | 12% | |
| $NH_2$-terminus | Glx | Ala | Ala | Phe | n.d. | Ala | n.d. | Phe | n.d. | Phe | n.d. | n.d. | |

See legends to tables VII-I.1 and 2.
b) Met determined as Hse.

Table VII-II.3 Amino acids released by CPA, CPB and LAP from fragments of CN-1, CN-8 and CN-9.

| Peptide | Enzyme | Incubation time(h) | Amino acids released[a)] |
|---|---|---|---|
| CN-1 Th-1 | | | ...-Asp-Gln-Ala |
| | CPA | 18 | 1.0 |
| CN-1 Th-2 | | | Ala-Trp-Asn |
| | LAP | 3 | 1.0 0.5 0.6 |
| CN-1 T-3 | | | ...-Asp-Ala-Ile-Lys |
| | CPA/B | 6 | 0.7 1.0 |
| CN-8 T-1 | | | ...-Asn-Asp-Ile-Arg |
| | CPA/B | 6 | 0.6 1.0 |
| | | | d) d) |
| CN-8 Th-3 | | | Ile-Gln-Tyr-Asn-Gly |
| | LAP | 3 | 1.0 0.8 1.0 0.8 0.9 |
| CN-8 Th-3a | | | Asn-Gly |
| | LAP | 3 | 1.0 1.0 |
| CN-9 SP-2 | | | Pro-Glu-Met |
| | LAP | 18 | 1.0 0.9 0.9 |

See legends to table VII-I.10.
d) Asn and Gln were partially separated under the conditions used during amino acid analysis.

VII-II.3.2.2 Isolation and sequence determination of SP-peptides

From the digest 2 fragments were isolated, together comprising all residues of CN-8. The N-terminal tripeptide SP-1 was used for amide-assignment (VII-II.3.4.3).

VII-II.3.2.3 Isolation and sequence determination of Th-peptides

From the digest 4 small peptides were isolated in pure form, and 2 fragments of these peptides. According to their composition Th-1 represented the first 3 residues of CN-8, whereas Th-3 and Th-4 together corresponded with the last 10 residues of T-4. Dipeptide Th-2 was identical to T-3. Th-4 and the fragment Th-4a were sequenced, which established the C-terminal sequence of CN-8 and CN-8 T-4.

VII-II.3.3 SEQUENCE ASSEMBLY OF PARTS OF CN-8

A survey is given in fig.VII-I.16. The N-terminal sequence of CN-8 was identified up to and including the 6th residue, and comprised the 5 residues of T-1. This 6th residue was Ala. Both T-2 and T-3 had Ala at the N-terminus. None of the other partial peptides of CN-8 gave a sequenced overlap between T-1 and T-2, T-2 and T-3 or T-3 and T-4. So, no alignment was possible with the information of CN-8. However, the sequences of T-2, T-3 and the first few residues of T-4, in that order, coincided with the sequence of the first 11 residues of Tm-8. Also the following non-sequenced residues of T-4 corresponded with the amino acids present in the next part of Tm-8. This information gave the proper alignment of the T-peptides of CN-8. T-2, T-3 and the sequenced residues of T-4 extended the sequence to position 17. The amino acid compositions of Th-3 and Th-3a indicated Gly at position 18. Fully structured Th-4 and Th-4a determined the order of residues 19 to 23. T-4 connected Th-3 to Th-4 by composition. CPA hydrolysis of CN-8 confirmed the last 2 residues of Th-4. The sequence thus obtained for the last 6 residues of CN-8 agreed with the corresponding sequence of Tm-8. (In fact Th-4 helped to complete Tm-8's sequence assembly). Methods used to characterize the amides present in CN-8 are summarized in table VII-VI.

VII-II.3.4 Experimental details

VII-II.3.4.1 Tryptic digestion of CN-8, fractionation and sequence determination of T-peptides. After conversion of N-terminal pyroGlu into Glu (III.1.12), digestion with trypsin (table III.1) and fractionation by HVE at pH 3.6 (2000 V, 75 min) gave pure T-1, T-2, T-3 and T-4. The amino acid compositions are given in table VII-II.2. T-1 migrated in 2 different fluorescamine-positive bands. Probably, the N-terminal pyroGlu was incompletely converted into Glu. Then, the more acidic spot should contain pyroGlu-T-1. The N-terminal sequence of 3 residues of this product was determined after conversion of pyroGlu to Glu. With CPB only Arg was released, CPA + CPB liberated both Ile and Arg (table VII-II.3). Of T-2, T-3 and T-4 the sequence of 4, 2 and 5 residues, respectively, was established.

VII-II.3.4.2 Digestion of CN-8 with S.protease, fractionation and sequence determination of SP-peptides. The digest obtained (III.1.7.3) was fractionated by HVE at pH 6.5 (3000 V, 40 min). Two peptides, one being ninhydrin-negative, were isolated in pure state as judged by amino acid analysis (table VII-II.2). SP-1, the N-terminal tripeptide with pyro-Glu at the N-terminal side as concluded from the absence of a derivative after dansylation, was used for amide assignment. According to Offord, SP-1 (composed of Glx, $Asx_2$) had a net charge of -2 showing beside the N-terminal pyro-Glu the presence of 1 Asp and 1 Asn. According to the specificity of the enzyme used for digestion, Asp was located at position 3 and Asn at position 2 in CN-8.

VII-II.3.4.3 Thermolytic digestion of CN-8, fractionation and sequence determination of Th-peptides. Digestion is described in III.1.7.4. Fractionation by HVE at pH 3.6 (2000 V, 75 min) gave pure Th-1, Th-2, Th-3 and Th-4 and the fragments Th-3a and Th-4a (table VII-II.2). Of Th-2, Th-4 and Th-4a the sequence was determined. Incubation of Th-3a with LAP showed the presence of Asn. LAP hydrolysis of Th-3 showed Gln to be present and confirmed the presence of Asn (see table VII-II.3).

VII-II.4 CN-9 (7 RESIDUES)

VII-II.4.1 SEQUENCE DETERMINATION OF CN-9a AND CN-9

Of CN-9a, assumed to be the C-terminal tripeptide of CN-9, the whole sequence was established, while of CN-9 the identity of 4 residues was determined (see fig.VII-I.16). Many problems were encountered during sequence analysis of CN-9. No residue could be detected in position 2, while repeated attempts to sequence beyond the 5th residue all failed. With CPA only Hse was released. If CN-9a was positioned correctly, the

residue in position 6 should be Glx. This was confirmed by the isolation of a dipeptide consisting of Pro and Glu after partial acid hydrolysis of CN-9 (0.03 N HCl for 8 h at 110°C) and the isolation of an SP-peptide, CN-9 SP-2, comprising Pro(1.07) Glu(0.99) Met(0.92) with N-terminal Pro (both were purified by HVE at pH 3.6 (4000 V, 50 min). Trp was identified in CN-9 on paper with a specific reagent (III.1.9.2.2) and should occupy position 2. Glx and Asx at positions 3 and 4 were identified as PTH-Glu and PTH-Asp, respectively. Bond splitting with S.protease after Asp 4 confirmed the acidic character. Glu in position 6 was identified after hydrolysis of CN-9 SP-2 with LAP. Combination of the results gave the structure for CN-9 as shown in fig.VII-I.16 which corresponded with the sequence found for the corresponding part of Tm-8. CPA and LAP results are given in chapter IV (table VI.4), while methods used to assign the amides are summarized in table VII-VI.

## VII-II.5 CN-10 (10 RESIDUES)

### VII-II.5.1 SEQUENCE DETERMINATION OF CN-10

The order of all residues was determined by dansyl-Edman degradation; incubation with LAP showed the presence of 1 Glu, 1 Asn and 1 Asp (table VI.4). This elucidated the sequence of CN-10 as given in fig. VII.I.16, but for the positions of Asp and Asn. This ambiguity was later resolved by studies on [Tm-8b] (VII-V.2 and table VII-VI).

## VII-III SEQUENCE DETERMINATION OF FRAGMENTS OF "CN-"PEPTIDES

The results obtained with the CN-peptides of atropinesterase were not sufficient for the alignment of all 8 Tm-peptides. Unfortunately, only CN-8 contained an internal Arg-residue; it provided the overlap between Tm-7 and Tm-8. Tm-1 and Tm-2 were connected via the N-terminal sequence of atropinesterase, but the other 5 connections were still missing. The large CN-peptides were supposed to comprise the missing sequences. However, these could not be isolated as intact CN-peptides, but only as a mixture of "CN-"peptides (see VI-II and fig.VI.5). To

obtain the needed information, it was attempted to isolate Arg-containing fragments of the "CN-"peptides after digestions with various endopeptidases. This approach succeeded in that numerous useful peptides were obtained and could be sequenced, which made the alignment possible of almost all Tm-peptides and furthermore provided data for the ranking of T-peptides in Tm-6.

### VII-III.1 SECONDARY CLEAVAGES OF "CN-"PEPTIDES

"CN-"peptides of different CNBr-cleavage experiments (pools II, see fig.VI.5) were digested with trypsin, pepsin or S.protease. After fractionation by ion-exchange chromatography (figs.VII-III.1-3), followed by electrophoretic separation on paper, many reasonably sized, pure peptides were isolated.

The largest number of fragments was recovered from the peptic digest. Four of these p-peptides were used for establishing connections between the Tm-peptides, as described in the subsequent section, 7 others contributed to the determination of internal sequences of Tm-5 and Tm-6 (VII-III.3). About 10 more were isolated and characterized; these confirmed already well-established other sequences and are not described. From the SP-digest 4 useful fragments (sp-peptides) were recovered, whereas among the 5 t-peptides isolated only 1 was used for a final sequence elucidation.

The p-, sp- and t-peptides were given code-names based on their assumed origin, i.e. they were positioned in one of the non-isolated CN-peptides of atropinesterase which were numbered in the order of their occurrence in the intact protein, as judged from the established primary sequence (fig.VII-VI). These peptides should comprise residues 27-99 ("CN-2"), 112-178 ("CN-(4+5)"; the Met-Thr bond between "CN-4" and "CN-5" was only partially cleaved), 115-178 ("CN-5") and 179-228 ("CN-6"). The isolated fragments of the "CN-"peptides were numbered according to their assumed position in the peptide (CN-x"p-y).

#### VII-III.1.1 Experimental details

For digestions with trypsin and S.protease, about 10 mg peptide material (pool II; fig.VI.5) was used, for peptic digestion about 30 mg.

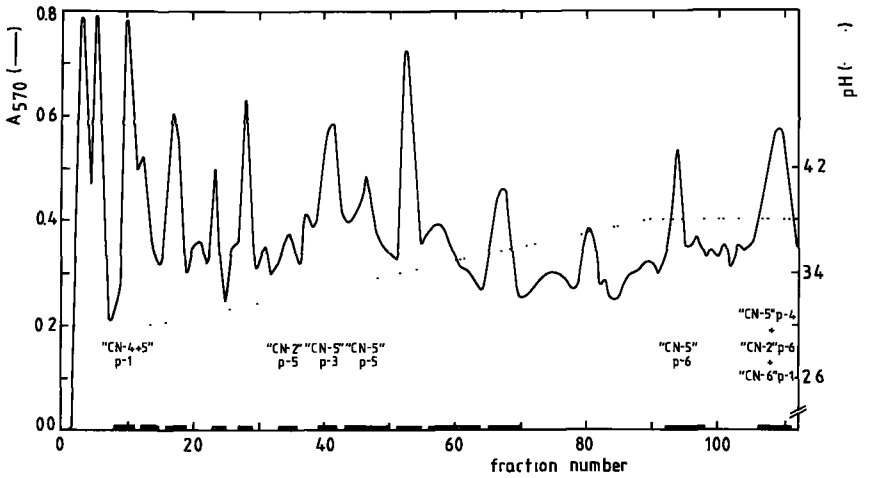



Fig.VII-III.1A. Fractionation of P-peptides of "CN-"fragments. A column of Aminex A-5 (26 x 0.9 cm) was eluted at 50°C with 50 ml 0.1 M pyridinium acetate, pH 3.0 (buffer A), followed by two steps gradient elution as described for fig. VII-I.2, at a flow rate of 49 ml/h. Buffer A: 175 ml; buffer B: 150 ml 0.5 M pyridinium acetate, pH 3.8; buffer C: 2.0 M pyridinium acetate, pH 5.0; elution was terminated with 2 M pyridine. Fractions of 3.9 ml were collected; 4% of the effluent was fed into the PA for detection of peptides (A570:——). Fractions were pooled as indicated. Pools with substantial amounts of p-peptides are marked as "CN-x"p-y. After fraction 115, a broad peak was pooled which was rechromatographed as shown in panel B.

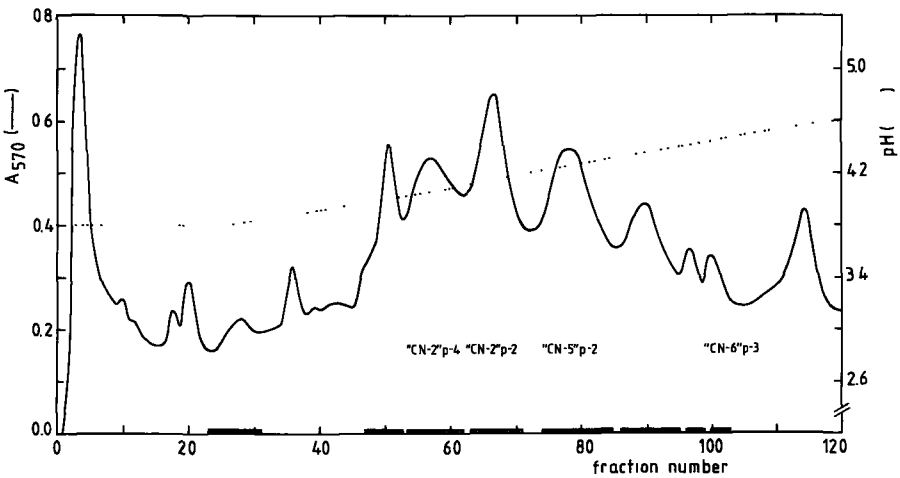



Fig.VII-III.1B. Chromatography of the pool of fractions after no.115 of panel A; elution was first with 50 ml buffer B, then a gradient formed with 250 ml B and 250 ml C.

Digestions were as described in III.1.7. Digests were fractionated on Aminex A-4 according to the procedures described in the legends to figs.VII-III.1-3. Chromatography of the peptic digest gave many peptide-containing peaks (fig.VII-III.1A/B). Amino acid analyses and end-group determination with dansylchloride indicated that none of the pools contained a pure peptide. A second fractionation by HVE at pH 3.6 (4000 V; 60 or 120 min) yielded a considerable number of pure peptides. Fractionation of the sp-digest resulted in a smaller number of identified fragments (fig.VII-III.2). Purification was by HVE at pH 6.5 (4000 V; 40 min). The separation of the t-peptides is shown in fig.VII-III.3; these were purified by HVE at pH 3.6.
Many peptides were characterized that were not needed for sequence establishments, and which -therefore- are only briefly mentioned. This concerns the following p-peptides: fragments of "CN-2"p-5 (residues 88-94, 89-94, 91-93) and p-6 (95-99), of "CN-(4+5)"p-1 (115-119), p-2 (120-130), p-3 (133-136, 133-143, 136-144), p-5 (152-163) and p-6 (163-166) and of "CN-6" p-3 (198-200). These were fully sequenced. Partially sequenced were "CN-2"p-1 (27-40) and "CN-2"p-3 (52-68). All p-peptides supposed to be formed from "CN-2" were identified. Fragments not formed among the p-peptides were the C-terminal 11 residues of "CN-5" and 2 regions of "CN-6" (11 and 27 residues). In addition to the t-peptide "CN-5"t-5 (VII-III.3), "CN-2"t-1 (27), t-2 (28-50; = Tm-2), t-3 (51-61), t-4 (62-64; 51-64 = Tm-3) and t-5 (65-71; = Tm-4)

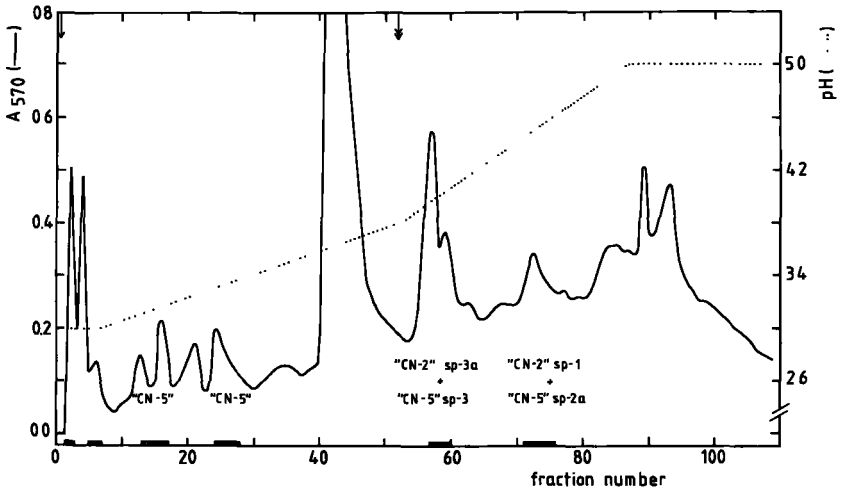



Fig.VII-III.2. Fractionation of the peptides resulting from digestion of "CN-"fragments with S.protease. A column of Aminex A-5 (26 x 0.9 cm) was eluted at 50°C with 15ml buffer A, followed by two steps gradient elution as described with fig.VII-I.2. at a flow rate of 49 ml/h. Conditions were as in fig.VII-III.1. First gradient: 100 ml buffer A; second: 75 ml buffer B. ↓, ↓ : start 1st and 2nd gradient. Of the effluent, 7% was fed into the PA (A570:——). Fractions were pooled as indicated.

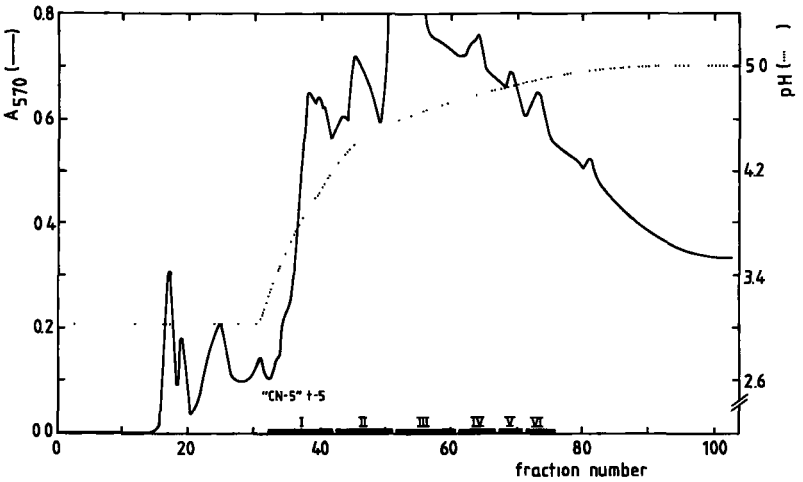



Fig.VII-III.3 Fractionation of the T-peptides of "CN-"fragments. A column of Aminex A-5 was used. All conditions were identical to those of fig.VII-III.2, except the elutions. Fifty ml of buffer A was followed by gradient going from 200 ml A to 200 ml C. Fractions were pooled as indicated.

were isolated. All sequences identified in these p- and t-peptides were in complete agreement with the earlier obtained data.

## VII-III.2 CHARACTERIZATION AND SEQUENCE DETERMINATION OF p- AND sp-PEPTIDES PROVIDING THE CONNECTIONS BETWEEN Tm-PEPTIDES

Four pure p-peptides, "CN-2"p-2 and p-4, "CN-5"p-2 and "CN-6"p-3 (fig.VII-III.1B), and 3 sp-peptides, "CN-2"sp-1, sp-3a and "CN-5"sp-2a (fig.VII-III.2), were isolated which overlapped with the ends of 2 Tm-peptides. Their amino acid compositions and the connections they established are shown in table VII-III.1. Of all peptides but one, the sequence was partially or completely determined by dansyl-Edman degradation, as is shown in figs.VII-III.6 and VII-VI. Additional information was obtained from digestions with CPA and/or CPB or LAP (table VII-III.4; fig.VII-III.6).

For complete sequence elucidation, "CN-2"p-4 (19 residues) had to be fragmented with trypsin. All 3 T-peptides were isolated (fig.VII-III.4) and characterized (table VII-III.2; fig.VII-III.6). The results fully established the junction between Tm-4 and Tm-5. "CN-5"p-2

Table VII-III.1 Amino acid composition of p- and sp-peptides of "CN-"fragments.

| Amino acid | "CN-2" sp-1 | "CN-2" p-2 | "CN-2"[b] sp-3a | "CN-2" p-4 | "CN-5" p-2 | "CN-5"[b] sp-2a | "CN-6" p-3 |
|---|---|---|---|---|---|---|---|
| Asx | | 1.16(1) | 0.37 | 2.00(2) | 1.11(1) | 0.30 | |
| Thr | | 1.93(2) | 0.92(1) | 1.08(1) | | | |
| Ser c) | | 1.00(1) | 2.56(2-3) | 2.09(2) | 1.95(2) | 1.92(2) | |
| Glx | 1.03(1) | | 0.31 | | | 0.36 | |
| Pro | | | | 1.03(1) | 1.07(1) | | |
| Gly | | 1.12(1) | 0.40 | 2.91(3) | | 0.52 | 1.01(1) |
| Ala | | 1.15(1) | 1.27(1) | 0.30 | 1.22(1) | 0.40 | 1.05(1) |
| Val | 0.95(1) | | | 0.95(1) | 1.96(2) | 1.86(2) | |
| Met | | | | | | | |
| Ile | | | | 0.96(1) | 0.82(1) | 0.99(1) | |
| Leu | | 1.89(2) | 0.23 | 1.20(1) | 1.89(2) | 2.06(2) | |
| Tyr | 0.95(1) | 0.78(1) | | 1.68(2) | | | |
| Phe | | | 0.99(1) | | | | 0.97(1) |
| His | | 0.97(1) | | 1.10(1) | 0.80(1) | | |
| Lys | | | 0.33 | 1.01(1) | 1.11(1) | 1.05(1) | |
| Arg | 1.06(1) | 1.00(1) | 1.01(1) | 1.01(1) | 1.07(1) | 1.11(1) | 0.97(1) |
| Trp | | | | | | | |
| Cys d) | | | | 1.93(2) | | | |
| No. of residues | 4 | 11 | 6-7[c] | 19 | 13 | 9 | 4 |
| $NH_2$-terminus | Arg | Leu | Thr | Asp | Ala | n.f. | Ala |
| Junction | Tm-1/Tm-2 | Tm-2/Tm-3 | Tm-2/Tm-3 | Tm-4/Tm-5 | Tm-5/Tm-6 | Tm-5/Tm-6 | Tm-6/Tm-7 |
| Residue number e) | 27-30 | 41-51 | 48-53/54 | 69-87 | 120-132 | 126-134 | 198-201 |

See also legends to table VII-I.1.
b) Peptides containing impurities.
c) The exact amount of Ser in "CN-2" sp-3a could not be determined. The interme diate value of 2.5 could be due to partial destruction during hydrolysis or the presence of a contaminating peptide.
d) Cys was determined as CMCys.
e) Numbers refer to the positions in atropinesterase.

(15 residues), too, was digested with trypsin; the 3 fragments were isolated (fig.VII-III.5) and analysed (table VII-III.2). Although only T-3 could be sequenced (fig.VII-III.6), composition and sequences data of "CN-5"p-2 and its T-peptides, together with the sequence information on Tm-5 and Tm-6 and the composition found for "CN-5"sp-2a allowed an unambiguous conclusion with regard to the ranking of these 2 Tm-peptides.

VII-III.2.1 Experimental details

All procedures were as described in chapter III. Details on the chromatography are given with the figures. Yields of the peptides with respect to CNBr-digested atropinesterase could not been given because different pools II had been combined after various quantities had been used for other experiments. CPA digestion of "CN-2"p-4 identified Asx (87) as Asp.

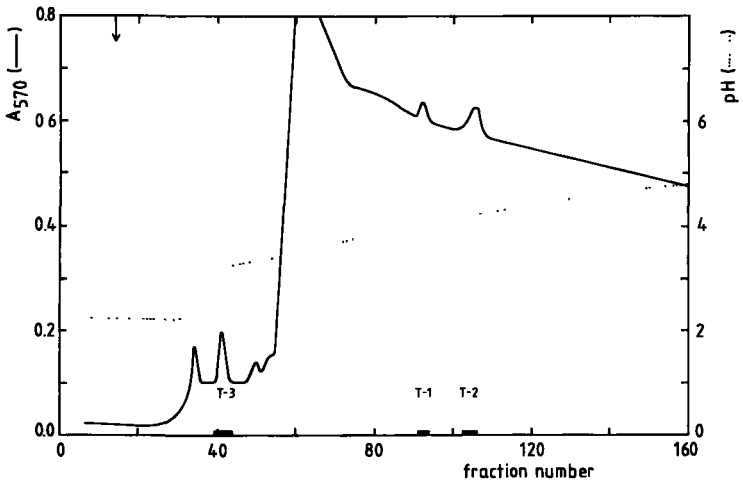



Fig.VII-III.4. Fractionation of T-peptides of carboxymethylated "CN-2"p-4. A column of Aminex A-5 (20 x 0.9 cm) was eluted at 50°C with 0.05 M pyridinium acetate, pH 2.4 (buffer A), at a flow rate of 26 ml/h, followed by a linear gradient elution (↓), going from 360 ml buffer A to 360 ml 2.0 M pyridinium acetate, pH 5.0. Fractions of 2.3 ml were collected; 14% of the effluent was fed into the PA for on line detection of peptides (A570:——). Fractions were pooled as indicated.

Table VII-III.2 Amino acid composition of T-peptides of "CN-2"p-4 and "CN-5"p-2a.

| Amino acid | "CN-2" p-4 T-1 | "CN-2" p-4 T-2 | "CN-2"[b] p-4 T-3 | "CN-5" p-2 T-1 | "CN-5" p-2 T-2 | "CN-5" p-2 T-3 |
|---|---|---|---|---|---|---|
| Asx | 1.00(1) | | 1.16(1) | 1.13(1) | | |
| Thr | | 1.09(1) | 0.48 | | | |
| Ser | | 1.18(1) | 1.29(1) | 1.23(1) | 1.03(1) | 0.39 |
| Glx | | | | | | |
| Pro | | 1.03(1) | | 1.10(1) | | |
| Gly | | 2.87(3) | 0.21 | 0.31 | | 0.36 |
| Ala | | | 0.29 | 1.23(1) | | |
| Val | | | 0.86(1) | | 0.97(1) | 0.97(1) |
| Met | | | | | | |
| Ile | | 0.99(1) | | 0.69(1) | | |
| Leu | 1.06(1) | | | | | 2.03(2) |
| Tyr | | | 1.76(2) | | | |
| Phe | | | | | | |
| His | | 0.91(1) | | 0.73(1) | | |
| Lys | | 0.94(1) | | 0.94(1) | | |
| Arg | 0.95(1) | | | | 1.00(1) | |
| Trp | | | | | | |
| Cys d) | | | 1.94(2) | | | |
| No. of residues | 3 | 9 | 7 | 7 | 3 | 3 |
| $NH_2$-terminus | | Gly | CM-Cys | | | Leu |

See legends to table VII-III.1.

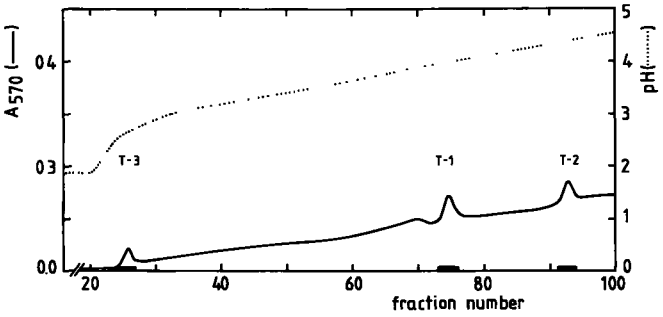



Fig.VII-III.5. Fractionation of T-peptides of "CN-5"p-2. A column of Aminex A-5 (20 x 0.9 cm) was eluted at 50°C, at a flow rate of 15 ml/h. Linear gradient elution was applied, going from 150 ml 0.05 M pyridinium acetate, pH 1.9, to 150 ml 2.0 M pyridinium acetate, pH 5.0. Fractions of 1.9 ml were collected; 21% of the effluent was consumed by on line detection of peptides in the PA (A570:——). Fractions were pooled as indicated.

## VII-III.3 CHARACTERIZATION AND SEQUENCE DETERMINATION OF p-, sp- AND t-PEPTIDES ESTABLISHING INTERNAL SEQUENCES IN Tm-5 AND Tm-6

For the alignment of the T-peptides of Tm-6, 4 of the "CN-"p-peptides provided the necessary overlaps, which in 2 cases were confirmed by sp-peptides. The amino acid compositions of these peptides, and the overlaps they provided, are shown in table VII-III.3. Three other "CN-5" fragments are included in this table, i.e. p-4, sp-3 and t-4, which gave confirmation of internal sequences in Tm-6 T-2 and T-3. The sequences determined in these fragments and their contribution to the assembly of Tm-6 are surveyed in fig.VII-I.8. Gln (152), the 4th residue in "CN-5"sp-4 was identified as the PTH-amino acid. This result was confirmed by LAP-digestion of "CN-5"p-5, which released 4 residues including Gln (152) and Asp (154) (table VII-III.4).

Three p-fragments of "CN-"peptides were isolated which confirmed sequences in Tm-5("CN-2"p-5 and p-6 and "CN-(4+5)"p-1, see table VII-III.3). CPA-digestion of "CN-2"p-5 released 4 residues and gave conclusive evidence for the Asp in position 94 in atropinesterase (see tables VII-III.4 and VII-VI).

Table VII-III.3 Amino acid composition of p-, sp- and t-peptides of "CN-"fragments of pool II (fig.VI.5) confirming parts of Tm-5, overlapping Tm-6 T-peptides or confirming parts of Tm-6.

| Amino acid | "CN-2" p-5 | "CN-2" p-6 | "CN-4+5" p-1 | "CN-5" p-3 | "CN-5" sp-2b | "CN-5" p-4 | "CN-5" sp-3 | "CN-5" sp-4 | "CN-5" p-5 | "CN-5" t-5 | "CN-5 p-6 | "CN-6" p-1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | 2.24(2) | 1.04(1) | | | 0.25 | 1.21(1) | 1.31(1) | 3.00(3) | 2.79(3) | 2.83(3) | | 0.93(1) |
| Thr | | | 0.92(1) | 1.82(2) | 1.83(2) | 1.06(1) | | 0.94(1) | | | | 0.92(1) |
| Ser | 0.90(1) | | 0.90(1) | 0.87(1) | | 0.22 | | 1.04(1) | 0.98(1) | 0.95(1) | | 0.21 |
| Glx | 0.95(1) | | | 1.09(1) | 1.13(1) | | | 2.05(2) | 2.05(2) | 1.28(1) | 0.98(1) | 1.88(2) |
| Pro | | | | 0.94(1) | 0.96(1) | | | 1.89(2) | 1.95(2) | 1.89(2) | | |
| Gly | | | 1.99(2) | 1.05(1) | 1.07(1) | 0.24 | 0.20 | 0.32 | | 0.28 | | |
| Ala | 1.03(1) | | 1.11(1) | 1.02(1) | 1.17(1) | | | | 0.23 | 1.13(1) | 1.03(1) | 1.10(1) |
| Val | 1.01(1) | | 1.10(1) | 1.21(1) | 0.81(1) | 1.81(2) | 0.90(1) | 1.10(1) | | | | |
| Met f) | | 0.78(1) | 0.54(1) | | | | | | | | | |
| Ile | | 1.10(1) | | 0.86(1) | | | | | | | | |
| Leu | | | 0.97(1) | 2.17(2) | 1.98(2) | 1.03(1) | | 2.01(2) | 1.14(1) | 0.97(1) | | 1.04(1) |
| Tyr | | | | | | 0.85(1) | 0.79(1) | | | | | |
| Phe | 1.90(2) | | | | | | | 0.98(1) | 1.09(1) | 1.86(2) | | |
| His | | | | | | | | | | | | |
| Lys | | 1.07(1) | | 0.90(1) | 1.06(1) | | | 0.99(1) | 0.91(1) | 1.09(1) | 0.99(1) | 2.13(2) |
| Arg | | | | | | | | | | | | |
| Trp | | | | | | n.d.(1) | n.d.(1) | | | | | |
| No. of residues | 8 | 4 | 8 | 12 | 10 | 7 | 4 | 14 | 11 | 12 | 3 | 8 |
| $NH_2$-terminus | Phe | Ile | Gly | Ile | Thr | ? | ? | Thr | Gln | Asp | Ala | Ala |
| Peptide of | Tm-5 | Tm-5 | Tm-5 | Tm-6 | Tm-6 | Tm-6 | Tm-6 | Tm-6 | Tm-6 | Tm-6 | Tm-6 | Tm-6 |
| Junction | | | | T-1/T-2 | T-1/T-2 | | | T-2/T-3 | T-2/T-3 | | T-3/T-4 | T-5/T-6 |
| Residue number e) | 88-95 | 96-99 | 112-119 | 133-144 | 135-144 | 145-151 | 145-148 | 149-162 | 152-162 | 154-165 | 164-166 | 174-186 |

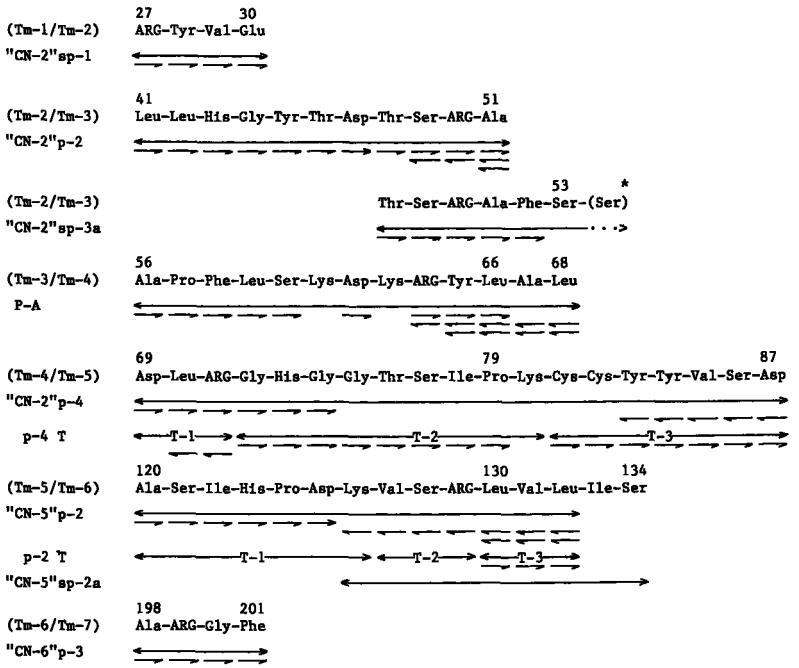
See legends to table VII-III.1.
f) Met determined as Hse.

Table VII-III.4 Amino acids released from "CN-"p- and t-peptides after incubation with CPA, CPA/CPB or LAP.

| Peptide | Enzyme | Incubation time (h) | Amino acids released[a)] |
|---|---|---|---|
| "CN-2"p-2 | | | Leu-Leu-His-Gly-Tyr-Thr-Asp-Thr-Ser-Arg-Ala |
| | CPA/B | 18 | 0.4 0.1 0.2 0.2 0.2 0.4 0.8 1.0 1.0 |
| | CPA | 6 | 1.0 |
| "CN-2"p-4 | | | ...-Pro-Lys-Cys-Cys-Tyr-Tyr-Val-Ser-Asp |
| | CPA | 18 | 2.0 1.1 1.0 1.0 |
| p-4 T-1 | | | Asp-Leu-Arg |
| | CPA/B | 18 | 0.4 1.0 |
| "CN-2"p-5 | | | Phe-Ala-Glu-Asp-Val-Ser-Asp-Phe |
| | CPA | 18 | 1.1 1.1 1.2 1.0 |
| "CN-5"p-2 | | | ...-Pro-Asp-Lys-Val-Ser-Arg-Leu-Val-Leu |
| | CPA/B | 18 | 0.9 1.1 0.9 1.9 2.0 |
| | CPA | 18 | 1.0 2.0 |
| "CN-5"p-5 | | | Gln-Lys-Asp-Phe-Pro-Leu-Asp-Asp-Pro-Ser-Glu |
| | LAP | 6 | 1.0 1.0 1.3 0.6 |
| "CN-5"t-5 | | | ...-Pro-Leu-Asp-Asp-Pro-Ser-Glu-Phe-Ala-Lys |
| | CPA/B | 18 | 1.0 1.0 1.0 |

See legends to table VII-I.10.

Fig.VII-III.6. Amino acid sequences of "CN-"p- and "CN"-sp-peptides and of P-A, a peptic peptide of atropinesterase connecting the Tm-peptides of the protein.




*) "CN-2"sp-3a was produced by aspecific bond splitting by S.protease at the C-terminal side of Ser.
. . . Presence of third Ser could not be established with certainty.

## VII-IV SEQUENCE DETERMINATION OF TWO PEPTIC PEPTIDES OF ATROPINESTERASE

Among a larger number of peptides derived from atropinesterase by digestion with pepsin (P-peptides) that contributed to the final elucidation of the primary structure of the protein, 2 were of particular importance. One, P-A, provided the overlap between 2 Tm-peptides, the other, P-B, helped to establish the internal sequence of Tm-8.

### VII-IV.1 SEQUENCE DETERMINATION OF P-A (13 RESIDUES)

The amino acid composition (a.o. 2 Lys + 1 Arg; see table VI.5) suggested that this fragment would comprise the C-terminal part of Tm-3.

Sequence determination of the residues 1-5, 7 and 9-11 confirmed this hypothesis; up to Arg (9) the identified residues were identical to the corresponding amino acids in Tm-3. The 10th and 11th residues were Tyr-Leu, followed by (Ala,Leu) according to the composition of P-A. This C-terminal sequence was confirmed by CPA and CPB digestions. CPA released Tyr (1.0), Leu (2.0) and Ala (1.1), whereas CPA + CPB gave one additional residue (0.9 Arg). Tyr-Leu (Ala,Leu) is the unique N-terminal sequence of Tm-4. A survey of these results is given in figs.VII-III.6 and VII-VI.

### VII-IV.2 SEQUENCE DETERMINATION OF P-B (11 RESIDUES)

The amino acid composition (table VI.5) was identical to that of Tm-8 P-2 (table VII-I.9) and also the N-terminal residue (Ile) and the amino acids released with CPA were the same. The identity of the first 9 residues was established by dansyl-Edman degradation. Confirmatory information was obtained by CPA digestion for different times (table VII-I.10).

## VII-V ISOLATION, PURIFICATION AND SEQUENCE DETERMINATION OF [Tm-]PEPTIDES, FRAGMENTS OF Tm-PEPTIDES GENERATED BY ASPECIFIC CLEAVAGES

In the early stages of this study, the optimal conditions for maleylation of atropinesterase and subsequent digestion with trypsin had not been worked-out yet, with the result that to a considerable extent aspecific cleavages occurred (see also chapter V). Probably, these cleavages were partly due to a particular vulnerability of certain peptide bonds in the maleylated protein (which manifested itself for instance after freezing and thawing of a solution), partly to chymotryptic-like action of the trypsin preparations. As a consequence, many fragments were formed of the Tm-peptides aimed at. A large number of these aspecific fragments ([Tm-]peptides) have been isolated and characterized. In later stages, several of the [Tm-]peptides appeared to be very useful as their sequences filled-in gaps in the sequences of Tm-peptides or gave a desired confirmation of some data.

## VII-V.1 ISOLATION AND PURIFICATION OF [Tm-]PEPTIDES

In principle, the procedures were identical to those used for the production and isolation of Tm-peptides (see chapter VI). Due to the formation of many small fragments, however, fractionation of the digest on Sephadex G-50 sf gave a different elution profile and, consequently, fractions were pooled differently (compare figs.VI.1 and VII-V.1). Further fractionation of the pools by ion-exchange chromatography (figs.VII-V.2-4), usually followed by paperelectrophoresis, yielded several peptides. The ones providing profitable information have been listed in table VII-V.1, where amino acid compositions and the assumed positions in atropinesterase as well as in the parent Tm-peptides are given.

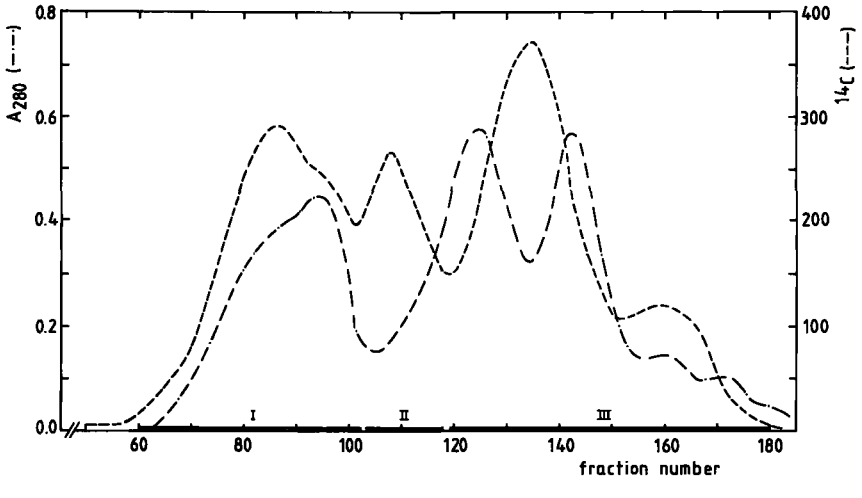



Fig.VII-V.1. Fractionation of aspecifically formed tryptic [Tm-]peptides of 1,4-[$^{14}$C]-maleylated atropinesterase on Sephadex G-50 sf. The column (136 x 2.3 cm) was eluted with 0.01 M Tris.HCl, pH 8.0, at 23 ml/h. Fractions of 2.3 ml were monitored at 280 nm (A280:—·—). Aliquots (0.7%) were removed for LSC(———). Fractions were pooled as indicated.

Table VII-V.1 Amino acid composition of [Tm]-peptides generated by aspecific bond splitting.

| Amino acid | [Tm-1a] * | [Tm-1b] ** | [Tm-2a] * | [Tm-3a] * | [Tm-3b] * | [Tm-5a] b) * | [Tm-6a] * | [Tm-6b] * | [Tm-7a] * | [Tm-7b] * | [Tm-8a] * | [Tm-8b] b) *** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Asx | 1.14(1) | 3.04(3) | 2.15(2) | 0.98(1) | 1.07(1) | 1.21(1) | 0.22 | | | 3.80(4) | 0.21 | 3.31(3) |
| Thr | | | | | | | 0.93(1) | | 1.82(2) | 1.08(1) | | 0.29 |
| Ser | 0.30 | 2.16(2) | 0.97(1) | 0.93(1) | 0.97(1) | 1.91(2) | | 0.84(1) | | 1.25(1) | 0.30 | 0.49 |
| Glx | 1.78(2) | 0.86(1) | 1.06(1) | | | | 1.05(1) | | | 3.61(4) | | 2.88(3) |
| Pro | 1.87(2) | | 0.81(1) | 0.95(1) | | 0.99(1) | 1.05(1) | | | 1.15(1) | 1.12(1) | 0.99(1) |
| Gly | 0.32 | 0.38 | 2.02(2) | | | 0.56 | 1.01(1) | | | 1.21(1) | 0.27 | 0.33 |
| Ala | 2.02(2) | 2.39(2) | 0.38 | 1.22(1) | | 1.13(1) | 0.33 | 2.20(2) | 1.09(1) | 0.31 | 2.80(3) | 1.41(1) |
| Val | 1.19(1) | 0.24 | 1.05(1) | | | 0.94(1) | 0.95(1) | | | | | 0.93(1) |
| Met | | 1.02(1) | | | | 0.25 | | | | 1.96(2) | | 1.17(1) |
| Ile c) | 1.52(2) | 2.58(3) | | | | 1.10(1) | | | | 1.82(2) | | 0.54 |
| Leu | 0.21 | 0.38 | 0.48 | 1.79(2) | 1.12(1) | 0.61 | 1.05(1) | | 1.97(2) | 1.22(1) | 1.09(1) | 1.93(2) |
| Tyr | | | 1.06(1) | | | | | | | | | |
| Phe | | | | 0.93(1) | | | | | | | 1.00(1) | 1.53(2) |
| His | | | | | | 0.90(1) | | | | | | |
| Lys | | 3.00(3) | | 2.09(2) | 1.93(2) | 0.84(1) | 1.01(1) | | 1.11(1) | 0.37 | 1.98(2) | 1.86(2) |
| Arg | | 0.95(1) | | 1.12(1) | 0.92(1) | 0.98(1) | | 0.96(1) | | 0.91(1) | | |
| Trp | 0.61(1) | | 0.89(1) | | | | | | | n.d.(1) | | n.d.(1) |
| No. of residues | 11 | 16 | 10 | 10 | 6 | 10 | 7 | 4 | 6 | 19 | 8 | 17 |
| Yield d) | 46% (80%) | 41% | 20% | 22% | 60% | 30% | 31% | 13% | 44% (44%) | 35% (98%) | 42% (72%) | 20% (76%) |
| $NH_2$-terminus | Glx | Asn | Tyr | Leu | Leu | Ala | Lys | Ser | Leu | Ile | Ala | Phe |
| Residue number e) | 1-11 | 12-27 | 28-37 | 55-64 | 59-64 | 120-129 | 138-144 | 196-199 | 213-218 | 219-237 | 238-245 | 256-272 |
| Tm-peptide f) | $1\text{-}11^N$ | $12\text{-}27^C$ | $1\text{-}14^N$ | $5\text{-}14^C$ | $9\text{-}14^C$ | $49\text{-}58^C$ | 9-15 | $67\text{-}70^C$ | 14-19 | $20\text{-}38^C$ | $1\text{-}8^N$ | $19\text{-}35^C$ |

See legends to tables VII-I.1 and 2.
b) Peptides containing impurities.
c) Low value due to the presence of an Ile-Ile bond.
d) Yields refer to percentages found after the final purification step; figures in parentheses show yields after one purification step.
e) Residue numbers refer to the positions in atropinesterase.
f) Tm-peptide: numbers refer to the positions of the residues in the Tm-peptides; N and C indicate the N- or C-terminal fragment of a Tm-peptide.
* See fig.VII-V.3.
** See fig.VII-V.2.
*** See fig.VII-V.4.

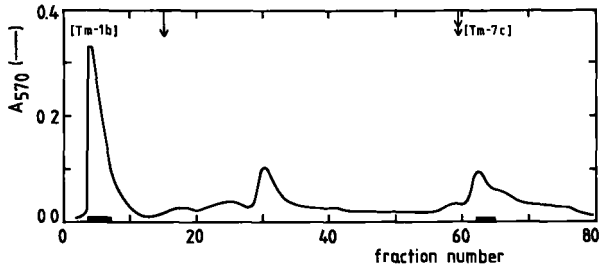



Fig.VII-V.2. Fractionation of pool II of fig.VII-V.1, containing aspecifically formed [Tm-]peptides. A column of AG 1-X2 (32 x 0.9 cm) was eluted at 40°C with 6 ml 1% collidinium acetate, pH 8.3, followed by 10 ml 1% collidine + 1% pyridine, brought with acetic acid to pH 7.5 (buffer A). Elution was continued by a two steps gradient, going from 40 ml buffer A to 40 ml 6% acetic acid and from 35 ml 6% to 35 ml 25% acetic acid, at a flow rate of 30 ml/h. Elution was ended with 18 ml 50% and 30 ml 70% acetic acid. Fractions of 1.9 ml were collected. Aliquots (1%) were used for detection of peptides with the PA (A570:——). Fractions were pooled as indicated.

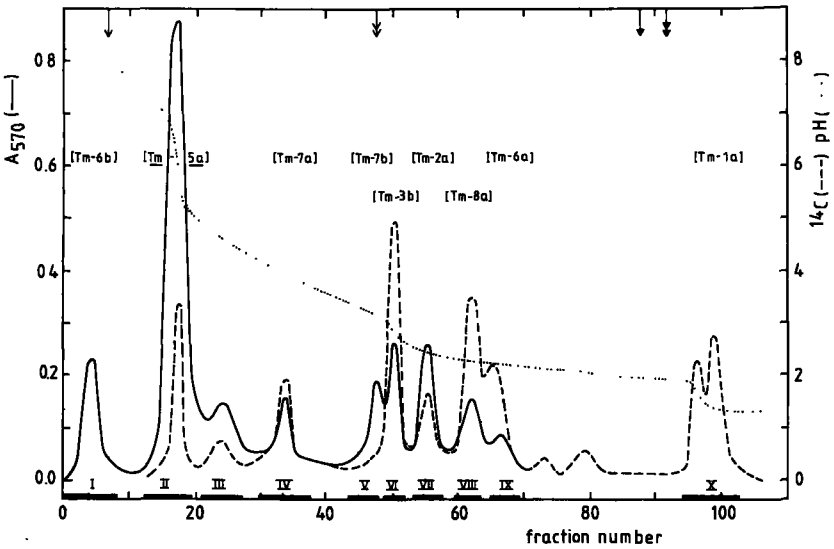



Fig.VII-V.3. Fractionation of pool III of fig.VII-V.1, containing aspecifically formed [Tm-]peptides. A column of AG 1-X2 (18 x 0.9 cm) was eluted at 40°C at a flow rate of 32 ml/h, as described with fig.VII-V.2; gradients going from 50 ml buffer A to 50 ml 6% acetic acid and from 50 ml 6% to 50 ml 25% acetic acid were used. Elution was finished with 10 ml 50% and 30 ml 70% acetic acid. Fractions of 2.1 ml were collected. Aliquots (1%) were used for detection of peptides in the PA (A570:——) and aliquots (1%) were used for LSC (---). Fractions were pooled as indicated. ↓, ↓ : start 1st and 2nd gradient; ↓, ↓ : start 50% and 70% acetic acid.

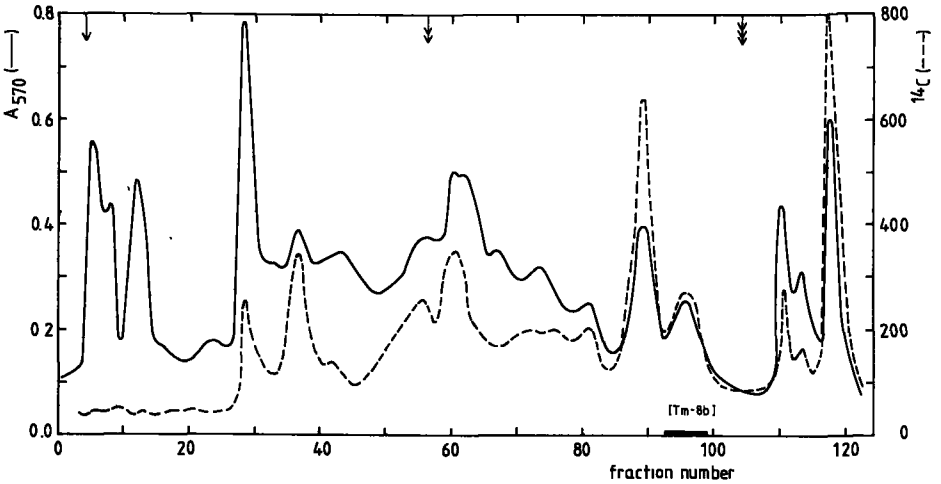



Fig.VII-V.4. Fractionation of pool I of fig.VII-V.1, containing aspecifically produced [Tm-]peptides. A column of DEAE-Sephadex A-25 (29 x 1.5 cm), was eluted at a flow rate of 53 ml/h, with 15 ml 2% 2-picoline + 1% pyridine, brought with acetic acid at pH 7.1 (buffer A), elution was continued with a two steps gradient going from 100 ml buffer A to 100 ml 6% acetic acid and from 100 ml 6% to 100 ml 25% acetic acid. Elution was ended with 10 ml 50% and 50 ml 70% acetic acid. Fractions of 3.5 ml were collected. Of the effluent 4% was used for on line detection of peptides with the PA (A570:——). Aliquots (1.4%) were used for LSC (---). ↓, ↓ : 1st, 2nd gradient; ↓ : 70% acetic acid.

VII-V.1.1 Experimental details

Maleylation and digestions with trypsin were as described in III.1.3 and 6. In a first experiment gelfiltration of the digest gave the elution profile shown in fig.VII-V.1. Desalting of pools II and III on Sephadex G-10 (63 x 2.0 cm) in 1% collidinium acetate, pH 8.3, and chromatography of the ninhydrin-positive material on Dowex 1-X2 (AG 1-X2) resulted in the isolation of the indicated [Tm-]peptides (figs. VII-V.2/3). Most peptides were pure enough for sequence studies, however, to obtain well-fitting amino acid compositions, usually additional purification was needed. Of the pools of fig.VII-V.3 demaleylated [Tm-1a], [Tm-6a], [Tm-6b], [Tm-7a] and [Tm-8a] were purified by HVE at pH 3.6 (2000 V, 90 min) and demaleylated [Tm-7b] and [Tm-2a] by fractionation on Sephadex G-15 in 5% acetic acid, while [Tm-5a] was rechromatographed on Dowex 1-X2 by using a gradient consisting of 150 ml 0.3 M picolinium-0.4 M pyridinium acetate, pH 7.5, to 150 ml 0.4 N acetic acid, which lowered the pH gradually from 7.5 to 6.0. In a second experiment the elution profile of the tryptic digest on Sephadex G-50 sf (not given) showed less difference with fig.VI.1, but CPA hydrolysis still indicated the presence of many C-terminal amino acids differing from Arg. Chromatography of fractions corresponding to pool III of fig.VI.1 on DEAE-Sephadex (fig.VII-V.4) showed the presence of many fragments. The pool indicated contained [Tm-8b] together with some contaminating amino acids (table VII-V.1). Desalting of fractions corresponding to pool b+IV of fig.VI.1 on Sephadex G-10 in 1% collidinium acetate, pH 8.3, and fractionation of the demaleylated peptide material by HVE at pH 3.6 (3000 V, 55 min) yielded [Tm-3a] (table VII-V.1).

## VII-V.2 SEQUENCE DETERMINATION OF [Tm-]PEPTIDES

The sequences of [Tm-3a], [Tm-7a] and [Tm-8a] provided essential or desired information on the structure of Tm-peptides, the other 9 [Tm-]peptides gave useful confirmations.

[Tm-3a]. The amino acid residues at 5 positions were identified, as indicated in fig.VII-I.3. Positions 3 and 4 were skipped because of the low amount of peptide available, while Lys in position 7 was not found. The sequence results provided the connection between Tm-3 Ch-1 and Ch-2. LAP liberated only Leu, while with CPA and CPB the last 7 residues were released, confirming the presence of Asp. CPA and CPB results are shown in table VII-V.3.

[Tm-7a]. The maleylated peptide was completely sequenced, which aligned Tm-7 T-2 and T-3 as shown in fig.VII-I.13. With CPA the last 2 residues were confirmed, while LAP released all constituting residues (table VII-V.3).

Table VII-V.2 Amino acid composition of T- and Th-peptides of [Tm-1b] and of SP-peptides of [Tm-8b].

| Amino acid | [Tm-1b] Th-2 | [Tm-1b] Th-3 | [Tm-1b] T-4 | [Tm-8b] SP-1 | [Tm-8b] SP-2 | [Tm-8b] SP-3 | [Tm-8b] SP-4 | [Tm-8b] (SP-(2+3)) |
|---|---|---|---|---|---|---|---|---|
| Asp | 2.24(2) | | | 1.37(1) | 1.18(1) | 1.03(1) | | 2.03(2) |
| Thr | | | | | | | | |
| Ser | | 0.39 | 0.30 | 0.32 | | | | 0.29 |
| Glu | 0.81(1) | | | 2.11(2) | | 1.09(1) | | 1.17(1) |
| Pro | | | | 1.00(1) | | | | |
| Gly | | 0.23 | 0.24 | 0.23 | | 0.20 | | 0.27 |
| Ala | 1.15(1) | 0.31 | 0.23 | 0.25 | 1.03(1) | | | 1.20(1) |
| Val | | | | | 1.03(1) | | | 0.90(1) |
| Met | | 1.03(1) | 0.87(1) | | 0.74(1) | | | 0.66(1) |
| Ile | 1.80(2) | 0.75(1) | | | | | | |
| Leu | | | | | | 0.88(1) | 1.13(1) | 1.00(1) |
| Tyr | | | | | | | | |
| Phe b) | | | | 0.53(1) | | | 0.80(1) | |
| His | | | | | | | | |
| Lys | 0.21 | 1.25(1) | 0.39 | | 1.06(1) | | 1.04(1) | 1.07(1) |
| Arg | | 0.97(1) | 1.13(1) | | | | | |
| Trp | | | | n.d.(1) | | | | |
| No. of residues | 6 | 4 | 2 | 6 | 5 | 3 | 3 | (8) |
| Yield c) | 23% | 28% | 22% | 30% | 20% | 30% | 45% | 20% |
| $NH_2$-terminus | Ile | Ile | Met | Phe | Met | Leu | Phe | Met |

See legends to table VII-I.1 and 2.
b) N-terminal Phe low due to preceding reaction of peptide with ninhydrin on paper.
c) Yields are relative to the amount of atropinesterase degraded.

Table VII-V.3 Amino acids released by CPA, CPB and LAP from [Tm-] peptides.

| Peptide | Enzyme | Incubation time (h) | Amino acids released[a] |
|---|---|---|---|
| [Tm-1a] | | | ...-Asp-Gln-Ala-Ala-Trp |
| | CPA | 18 | 0.5 1.8 1.0 |
| [Tm-3a] | | | Leu-Ala-Pro-Phe-Leu-Ser-Lys-Asp-Lys-Arg |
| (dem.) | LAP | 18 | 1.0 |
| " | CPA/CPB | 18 | 0.9 1.0 1.1 1.0 2.0 n.d. |
| [Tm-3b] | | | Leu-Ser-Lys[d]-Asp-Lys[d]-Arg |
| (mal.) | LAP | 18 | 1.0 0.9 0.8 0.8 |
| [Tm-7a] | | | Leu-Thr-Ala-Lys[e]-Thr-Leu |
| (mal.) | CPA | 18 | 0.8 1.0 |
| (mal.) | LAP | 18 | 2.0 1.9 1.1 |

See legends to table VII-I.10; (mal.) refers to maleylated peptides, (dem.) refers to demaleylated peptides.
d) Lys co-eluted with $NH_3$ and could not be determined quantitatively.
e) ε-maleyl-Lys eluted at the position of Met and was determined only qualitatively as the standard value was unknown.

[Tm-8b]. This not totally pure C-terminal fragment of Tm-8 had to give the information needed to complete the sequence of Tm-8. The residues in positions 1, 3 and 4 were identified, but in position 2 no amino acid was found. Secondary cleavage with S.protease after demaleylation, and fractionation of the digest by paperelectrophoresis gave the 5 SP-peptides mentioned in table VII-V.2. Four peptides comprised all residues of [Tm-8b], the 5th, SP-(2+3), comprised SP-2 and SP-3 due to incomplete cleavage of the Asp-Leu bond. Of all SP-peptides the sequence was partially or completely elucidated as shown in fig.VII-I.16. SP-2 and SP-(2+3) connected Tm-8 P-3 and P-4 by sequence. In SP-1, being the N-terminal peptide of [Tm-8b] according to the N-terminal sequence, Trp was identified by specific colour reaction; this should occupy position 2 where no residue could be found in the sequence studies. The specificity of S.protease indicated the acidic character of the C-terminal residues of SP-1, SP-2 and SP-3. Asn in SP-3 was identified as PTH-derivative. Surprisingly, no cleavage had occurred at Glu in position 3 in [Tm-8b].

[Tm-1a]. This N-terminal fragment of atropinesterase was recognized by the presence of a $^{14}$C-maleyl group presumably at the $\alpha$-$NH_2$ in a peptide without Lys, a blocked N-terminus and by the sequence of the first 5 residues being identical to the N-terminal sequence of the protein (fig.VII-I.1). With CPA the last 3 residues were definitely identified. The preceding residue, an amide co-eluting with Ser, was tentatively recognized (table VII-V.3). According to Offord, demaleylated [Tm-1a] had a net charge of -2, indicating the presence of 2 acids and 1 amide, which was consistent with the assignment of amides in peptides of CN-1 (table VII-VI) and with sequencer results of the protein (chapter IV).

[Tm-1b]. The sequence of the first 12 residues was determined by dansyl-Edman degradation. After digestion of demaleylated [Tm-1b] with thermolysin and trypsin, respectively, and fractionation by paperelectrophoresis, 2 almost pure Th-peptides, Th-2 and Th-3, and 1 almost pure T-peptide, T-4, were obtained (see table VII-V.2). Sequence determination of the Arg-containing C-terminal peptides, Th-3 and T-4, elucidated the sequence of the last 4 residues of [Tm-1b]. These

sequence results and those of [Tm-1b] itself provided an independent determination of the C-terminal sequence of Tm-1, which fully confirmed the sequencer results and the analysis of CN-1 as is shown in fig.VII-I.1.

[Tm-2a] and [Tm-3b]. The structures found for these peptides were in complete agreement with those of the identical peptides Tm-2a and Tm-3 Ch-2, as shown in fig.VII-I.3.

[Tm-5a], [Tm-6a] and [Tm-6b]. These completely sequenced peptides confirmed the C-terminal structure of Tm-5, an internal sequence of Tm-6 (138-144) and the C-terminal sequence of Tm-6, respectively, as shown in fig.VII-I.8.

[Tm-7b]. Of this peptide comprising the last 19 residues of Tm-7 the first 3 residues were sequenced, which confirmed the N-terminal residue of Tm-7 SP-1 Ch-d as shown in fig.VII-I.13.

[Tm-8a]. This slighty contaminated, maleylated octapeptide was sequenced all through, which gave confirmation of the first 8 residues of Tm-8 as shown in fig.VII-I.16; the C-terminal Phe was confirmed by CPA hydrolysis.

VII-V.2.1 Experimental details

Digestions with S.protease, trypsin, thermolysin, CPA/B and LAP were as given in chapter III. CPA/B and LAP results are mentioned in table VII-V.3. Amide assignments have been summarized in table VII-VI. The SP-digest of [Tm-8b] was fractionated by HVE at pH 3.6, 4000 V, 45 min. The same applies to the Th- and T-digests of [Tm-1b]. The Th- and T-peptides of [Tm-1b] not mentioned in table VII-V.2 were tentatively identified in eluates.

## VII-VI DETERMINATION OF THE PRIMARY STRUCTURE OF ATROPINESTERASE

The amino acid sequence of the protein was determined unambiguously by alignment of the 8 Tm-peptides, which could be positioned in order with sequencer results of the protein (chapter IV), with p- and sp-peptides of "CN-" fragments (VII-III), with CN-8 (VII-II.3) and one P-peptide of the protein, P-A (VII-IV). This alignment is visualized in fig.VII-VI.

Sequencer results gave the overlapping sequence from Tm-1 to the 4th residue of Tm-2. The amino acid composition and sequence of "CN-2"-sp-1, which possessed only the C-terminal residue of Tm-1 at its N-terminal side, was consistent with the joint determined. Tm-2 and Tm-3 were connected by "CN-2"sp-3a, a junction which was confirmed by the amino acid sequence of "CN-2"p-2 which overlapped Tm-2 by 10 and Tm-3 by one residue. The connection between Tm-3 and Tm-4 was established by sequence, CPA and composition results of P-A, while Tm-4 and Tm-5 were overlapped by "CN-2"p-4, both by 3 sequenced residues. Tm-5 and Tm-6 were overlapped by composition of "CN-5"sp-2a and p-2; the N-terminal sequence of the latter coincided with a sequence of 6 residues of Tm-5, while the sequence of the last 3 residues of "CN-5"p-2, as determined in "CN-5"p-2 T-3, was identical to the N-terminal sequence of Tm-6. CPA and CPB hydrolysis results of "CN-5"p-2 confirmed the determined position of "CN-5"p-2 T-3. Tm-6 and Tm-7 were connected by sequence results of "CN-6"p-3. Finally, Tm-7 and Tm-8 were aligned by CN-8 and fragments of CN-8. Some of the established joints were confirmed by the composition of peptic peptides of atropinesterase.

Amino acid compositions of most of the peptides that overlap with Tm-peptides, are given in table VII-III.1. The amino acid compositions of P-A and CN-8 are given in tables VI.5 and 3, respectively.

The amino acid sequences of these peptides are given in fig.VII-III.6 but for the sequence of CN-8 which is given in fig.VII-I.16. CPA and CPB hydrolysis results of atropinesterase were consistent with the established C-terminal structure.

The primary structure of atropinesterase constructed from the fragments mentioned above is shown in fig.VII-VI; only those peptides which provided sequences overlapping with well-established sequences of Tm-peptides are shown. Evidence for the given sequences of Tm-peptides can be found in figs.VII-I.1, 3, 8, 13 and 16.

Amide assignment in atropinesterase. The many amides and acidic residues which are present in atropinesterase were characterized with various methods as shown in table VII-VI. All together 12 Asn and 7 Gln and 19 Asp and 12 Glu residues were found.

Table VII-VI Amide assignment of residues in atropinesterase (AE)

| Amino acid | position in Tm-1/CN-1 | position in AE | identified in | | Identification method | Conclusion |
|---|---|---|---|---|---|---|
| Glx a) | 1 | 1 | [Tm-1a] | | PTH | Glu |
| " | 1 | " | CN-1 | SP-1 | SP /HVE(m=0.90, -1) | " |
| " | 1 | " | AE | | PTH-HP | " |
| Asx a) | 7 | 7 | CN-1 | SP-2 | SP /HVE(m=0.31, -1) | Asp |
| " | 7 | " | AE | | PTH-HP | " |
| Glx a/b) | 8 | 8 | AE | | PTH-HP | Glu/Gln[c] |
| Asx b) | 12 | 12 | CN-1 | Th-2 | LAP | Asn |
| " | 12 | " | AE | | PTH-HP | " |
| Glx b) | 19 | 19 | AE | | PTH-HP | Glu/Gln[c] |
| Asx b) | 21 | 21 | AE | | PTH-HP | Asn |
| Asx b) | 22 | 22 | CN-1 | SP-3 | SP | Asp |
| " | 22 | " | AE | | PTH-HP | " |
| | Tm-2 | | | | | |
| Glx d) | 3 | 30 | Tm-2 | SP-1 | SP/LAP | Glu |
| " | 3 | " | AE | | PTH-HP | " |
| Asx d) | 6 | 33 | Tm-2 | SP-2 | PTH | Asn |
| Asx d) | 10 | 37 | Tm-2 | SP-2 | PTH | Asp e) |
| Asx | 20 | 47 | Tm-2 | | CPA/CPB | Asp |
| " | 20 | " | Tm-2 | P-2 | PTH | " |
| | Tm-3 | | | | | |
| Asx | 12 | 62 | Tm-3 | | LAP | Asp |
| " | 12 | " | Tm-3a | | CPA/B | " |
| | Tm-4 | | | | | |
| Asx | 5 | 69 | Tm-4 | | CPA/CPB | Asp |
| | Tm-5 | | | | | |
| Asx | 16 | 87 | Tm-5 | | PTH-HP | Asp |
| " | 16 | " | "CN-2" | p-4 | CPA | " |
| Glx | 19 | 90 | Tm-5 | P-3 | LAP | Glu |
| Asx | 20 | 91 | Tm-5 | P-3 | LAP | Asp |
| " | 23 | 94 | Tm-5 | P-3 | LAP | Asp |
| " | 23 | " | "CN-2" | p-5 | CPA | " |
| " | 26 | 97 | "CN-2" | p-6 | PTH | Asp |
| " | 32 | 103 | CN-3 | | LAP | Asn |
| " | 54 | 125 | Tm-5 | P-5 | PTH | Asp |
| " | 54 | " | "CN-5" | p-2 | PTH | " |
| | Tm-6 (Cm-6a) | | | | | |
| Glx f) | 15 | 144 | Cm-6a | T-2 | PTH | Glu |
| " | 15 | " | Tm-6 | SP-1 | SP | " |
| Asx f) | 19 | 148 | Tm-6 | P-3 | CPA | Asp |
| " | 19 | " | Tm-6 | SP-2 | SP | " |
| Glx f) | 23 | 152 | "CN-5" | p-5 | LAP | Gln |
| " | 23 | " | "CN-5" | sp-4 | PTH | " |
| Asx g) | 25 | 154 | Cm-6a | T-3 | PTH | Asp |
| " | 25 | " | "CN-5" | p-5 | LAP | " |
| Asx g) | 29 | 158 | Cm-6a | T-3 | PTH | Asp |
| Asx g) | 30 | 159 | " | " | PTH | Asp |
| Glx g) | 33 | 162 | " | " | PTH | Glu |
| " | 33 | " | Tm-6 | SP-3 | SP | " |
| Glx | 37 | 166 | Cm-6a | T-4 | PTH/LAP | Glu |
| " | 37 | " | Tm-6 | P-5 | CPA | " |
| " | 37 | " | Tm-6 | SP-4 | SP | " |
| Asx | 46 | 175 | Cm-6a | $T-5_1$ | PTH | Asp |
| Asx | 47 | 176 | Cm-6a | $T-5_2^1$ | PTH | Asp |
| " | 47 | " | " | $T-5^2$ | PTH | Asn |
| Asx | 52 | 181 | Cm-6a | T-6 | LAP | Asn |
| Glx | 56 | 185 | Cm-6a | T-7 | PTH | Glu |
| Glx | 57 | 186 | " | " | PTH | Glu |
| " | 57 | " | Tm-6 | SP-5 | SP | " |

*Table VII-VI Amide assignment of residues in atropinesterase (AE)*

| Amino acid | position in Tm-7 | position in CN-8 | identified in | Identification method | Conclusion |
|---|---|---|---|---|---|
| Asx | 6 | 205 | Tm-7 | PTH | Asn |
| Asx | 24 | 223 | Tm-7 NBS-3 | PTH | Asn |
| Glx | 25 | 224 | " " | PTH | Gln |
| Asx | 26 | 225 | " " | PTH | Asn |
| Glx | 27 | 226 | " " | PTH | Gln |
| Glx h) | 31 | 230 | Tm-7 SP-1 | SP | Glu |
| Glx h/k | 34 1 | 233 | CN-8 | pyro-Glu 1) | Gln |
| Asx h/k | 35 2 | 234 | CN-8 SP-1 | HVE 1) | Asn |
| Asx h/k | 36 3 | 235 | Tm-7 SP-2a | SP | Asp |
| " | 36 | " | CN-8 SP-1 | SP 1) | " |
| | Tm-8 | CN-8 | | | |
| Glx | 10 | 15 247 | CN-8 Th-3 | LAP | Gln |
| Asx | 12 | 17 249 | " " | LAP | Asn |
| Asx | 12 | 17 " | " Th-3a | LAP | " |
| Glx | 21 | 258 | CN-9 | PTH | Glu |
| Asx | 22 | 259 | " | PTH | Asp |
| Glx | 24 | 261 | CN-9 SP-2 | SP/LAP | Glu |
| " | 24 | " | ⌈Tm-8b⌋ SP-1 | SP | " |
| Asx m) | 29 | 266 | " SP-2 | SP | Asp |
| Asx m) | 31 | 268 | " SP-3 | PTH | Asn |
| Glx m) | 32 | 269 | " " | SP | Glu |

Identification methods used: PTH, identification as PTH-amino acid after manual Edman degradation on TLC or HPLC (PTH-HP); SP, specificity of S.protease; HVE, high voltage paperelectrophoresis at pH 6.5 (The mobility m, was measured relative to Asp, charge was calculated by the method of Offord (93); CPA/CPB, LAP, identification as amino acid after digestion of the peptide with the exopeptidase.

a) HVE(m=0.40, -2) of ⌈Tm-1a⌋ showed the presence of 2 acidic residues and 1 amide.
b) HVE(m=0.21, +1) of CN-1 SP-3 showed the presence of 4 amides and 1 acidic residue beside 2 basic residues.
c) Taking into account the HVE results of ⌈Tm-1a⌋ and CN-1 SP-3, the additional evidence given for the character of residues on positions 1, 7, 12, 21 and 22 and the possibility of Gln to be converted to Glu, Glx-residues on positions 8 and 19 were assigned as amides.
d) HVE of Tm-2 P-1 (m=0.34, -2) was consistent with the presence of 2 acidic residues and 1 amide.
e) Aspecific splitting of Asp-Pro, causing the production of Tm-2a also showed the presence of Asp.
f) CPA/CPB hydrolysis of Cm-6a T-2(139-154) indicated the presence of 1 Asp, 1 Glu and 1 amide.
g) HVE(m=0.58, -3) of Cm-6a T-3(154-165) indicated the presence of 4 acidic beside 1 basic residue.
h) CPA and CPB results of Tm-7 and HVE results of NBS-3 (m=0.21, -1) were consistent with the presence of 1 Asp, 1 Glu and 2 amides behind Thr 30.
k) HVE of Tm-7 SP-2a (m=0.39, -1) showed the presence of 1 acid and 2 amides.
l) HVE of CN-8 SP-1 (m=1.16, -2) showed the presence of 1 acid and 1 amide beside a blocked N-terminus.
m) LAP results of Tm-8 P-4 and CN-10 and CPA/B results of Tm-8 confirmed the presence of 1 Asn, 1 Asp and 1 Glu at the C-terminal side of Tm-8.

## DISCUSSION

In general, it has been attempted to identify each residue in the constituting peptides at least twice, preferably via independent routes, if not by direct sequencing then by inference from compositions, CPA/B-results or LAP-data. For most positions in atropinesterase this approach succeeded. Whenever possible, sequenced overlaps of at least 2 residues in both directions were determined. Consequently, as a whole the endresult can be considered with confidence. In practice it is difficult, however, to obtain completely reliable results for all parts of a protein, and the experiences of others could urge to some caution; during the composition of the Atlas of Protein Structure and Sequence by Dayhoff over 10% of the published sequences were revised. Atropinesterase, like most sequenced proteins, appeared to comprise distinct parts that caused problems either in the isolation of sufficient amounts of pure fragments, or in the sequence analysis. For instance, isolation of the pure active serine Tm-peptide and certain fragments thereof proved extremely difficult, whereas 3 large CN-peptides could not be isolated at all because of poor solubility and a tendency to aggregate. Sequencing problems were encountered with Tm-3 and Tm-7. On the other hand, the sequence of the greater part of atropinesterase could be determined with great reliability because relevant fragments were obtained repeatedly in reasonable yields and could be sequenced without problems. Furthermore, many shorter or longer peptides have been isolated in the course of these investigations that were not described. None of these fragments conflicted with the primary structure as given in this chapter.

# CHAPTER VIII EVOLUTIONARY RELATIONSHIP BETWEEN ATROPINESTERASE AND OTHER SERINE-ENZYMES

In order to determine to which extent structural relationship exists between atropinesterase and other serine-enzymes, beside the functional relationship, firstly we compared the complete amino acid sequence of atropinesterase with the sequence of known serine-esterases and an essential region of the protein with segments of every sequence known. Secondly we tried to establish whether some similarity in spatial structure could be shown by comparing the predictions on the secondary structure of atropinesterase with the predicted or determined structure of a few other enzymes. Based on the results of these studies a conclusion was drawn with regard to a possible evolutionary relationship between atropinesterase and other serine-hydrolases.

## VIII.1 DETERMINATION OF SEQUENCE HOMOLOGY

Once the primary structure of atropinesterase was elucidated, we tried to determine the structural relationship between atropinesterase and other serine-enzymes, especially those having characteristics similar to the ones of atropinesterase, such as the capacity to hydrolyze ester bonds and the irreversible blocking of the active serine residue by organophosphorus compounds. Therefore, special attention was paid to the trypsins, chymotrypsins and subtilisins, which -in this respect- are related to atropinesterase and the structures of which have been thoroughly studied. Serine-hydrolases of this type have been classified by Dayhoff into two different superfamilies, i.e. the "prothrombin-related peptide hydrolases" comprising a.o. the trypsin- and chymotrypsin-family, and the "subtilisins" with the subtilisins as the only family (36, 138). This classification of proteins with known structures into superfamilies and the subdivision into families and subfamilies was based on detectable similarity between sequences. Families were defined as groups of proteins with sequences differing by less than 50%, while members of one subfamily show less than 20% difference in primary structure.

As a first step in our search for relationship between the newly sequenced protein and other serine-enzymes, we compared the amino acid sequence of atropinesterase in the immediate vicinity of the active serine with the active site pentapeptides of several serine-enzymes (35).

Secondly, we searched for indications of evolutionary relationship between an essential region of atropinesterase and known sequences in other proteins, by comparison of the active site sequence of atropinesterase, a 25 residues-segment with the active serine located at the center, with segments of the same length of every other sequence known. This study was performed with a computer method developed by Dayhoff (34).

Thirdly, the relationship between atropinesterase and other serine-hydrolases was investigated by comparison of the complete sequence of atropinesterase with the known sequences of a few serine-enzymes. This study was performed with another computer program by Dayhoff (34, 54) which opens the possibility to establish a common origin for 2 proteins even when in their sequences a large number of amino acid substitutions have accumulated. With this method we tried to detect any significant similarity between atropinesterase and the families of mammalian pancreatic serine-enzymes (the trypsins and chymotrypsins), by comparing the complete sequence of atropinesterase with the known structure of bovine trypsinogen (15, 62, 61, 69, 85, 129), one of the members of these related families. The amino acid sequence of atropinesterase was also compared with the structure of a member of another family belonging to the prothrombin-related enzymes, viz. the bacterial α-lytic protease from *Myxobacter* 495 (95, 19), which is only distantly related to the trypsin- and chymotrypsin-family, and -in the third place- with the sequence of subtilisin from Bacillus amyloliquefaciens (80, 94, 2), one of the three members of the second recognized superfamily of functionally related serine-hydrolases. Finally, the sequence of atropinesterase was compared with the sequence of 2 recently structured bacterial serine-enzymes, differing in function and other properties, viz. cephalosporinase (68) and alkaline phosphatase (71, 17) from Escherichia coli, in a search for any structural homology.

## VIII.1.1 RESULTS OF HOMOLOGY STUDIES

The active site pentapeptide of atropinesterase: Gly-His-Ser*-Met-Gly was compared with the known active site pentapeptides of other serine-enzymes, which could be divided in several classes (35) as is shown in table VIII.1. All prothrombin-related enzymes revealed a common sequence: Gly-Asp-Ser*-Gly-Gly. The sequence: Gly-Thr-Ser*-Met-Ala was found for the subtilisins and for some other microbial proteases. A third sequence, Gly-Glu-Ser*-Ala-Gly, was found in mammalian carboxylesterases and cholinesterases. The active serine pentapeptides of the functionally more remote alkaline phosphatases, phosphoglucomutases and cephalosporinase showed a further 3 sequences. The active site pentapeptide of atropinesterase did not show 100% sequence homology with any of the pentapeptides compared. However, it had beside the active serine itself, one residue in common with all "true" serine-hydrolases, viz. the Gly located two positions from the active serine at the N-terminal side. The second Gly, at the C-terminal side of the active site pentapeptide of atropinesterase was found in the prothrombin-related enzymes and also in the mammalian esterases, but not in the subtilisins. With the subtilisins, atropinesterase shared the Met next to the active serine. This comparison did not give any solid indication concerning sequence homology.

Table VIII.1 Active site pentapeptides of serine-enzymes

| | | |
|---|---|---|
| ATROPINESTERASE (Pseudomonas) | : Gly-His-Ser*-Met-Gly | |
| prothrombin-related enzymes | : Gly-Asp-Ser*-Gly-Gly | a) |
| subtilisins<br>alkaline proteases (A.Flavus) | : Gly-Thr-Ser*-Met-Ala | a) |
| carboxylesterases (mammalian)<br>serum cholinesterase (horse) | : Gly-Glu-Ser*-Ala-Gly | a) |
| other enzymes: | | |
| alkaline phosphatase (E.coli) | : Thr-Asp-Ser*-Ala-Ala | a) |
| phosphoglucomutases | : Thr-Ala-Ser*-His-Asp | a) |
| cephalosporinase (E.coli) | : Leu-Gly-Ser*-Val-Ser | b) |

a) for references see ref. 35 Dayhoff, M.O.
b) for reference see ref. 68 Jaurin, B. and Grundstrom, T.

Comparison of the extended active site region of atropinesterase with known sequences was performed by the computer program "SEARCH" of Dayhoff (34). A 25 residues-segment of atropinesterase, comprising residues 98-122 with the active serine located at the center (position 110), was compared with all 25 residues-segments of each of the sequences of 1809 proteins present in the database of Dayhoff at the end of April 1982 (37). For each segment-comparison a homology score was computed. These scores showed an approximately normal distribution, which indicated that atropinesterase in all probability did not have any evolutionary relationship with the sequences present in the data collection.

Next, the complete primary structure of atropinesterase was compared with the structure of the following serine-enzymes: bovine trypsinogen (229 residues) subtilisin from Bacillus amyloliquefaciens (275 residues) and α-lytic protease from Myxobacter 495 (198 residues), a bacterial member of the prothrombin-related enzymes. Furthermore, comparison was made with the primary structures of two serine-enzymes of Escherichia coli: cephalosporinase (377 residues), the structure of which was deduced from the nucleotide sequence, and with alkaline phosphatase (471 residues), the sequence of which was partially determined and partially deduced from the nucleotide sequence. (The established sequences of serine-enzymes are published by M.O. Dayhoff (34)). To determine whether similarity existed between atropinesterase and any of these proteins, the computer program "RELATE" was used (34, 54). In this program all segments of 25 residues present in atropinesterase were compared for sequence homology with all segments of the same length of the second enzyme. For each comparison of 2 segments a score was determined. Then, the probability was computed that the distribution of segment scores obtained from the comparison of the two real sequences, could have been derived from the distribution of scores obtained from the comparison of randomly permuted sequences (with the same amino acid compositions as the real sequences). In other words, the probability was computed that any similarity found is due to chance. The numerical value of this probability determines whether or not a newly determined sequence can be placed into an existing

Table VIII.2A: Segment comparison scores obtained after comparison of atropinesterase with other serine-enzymes with the program RELATE.

Table VIII.2B: Correlation between SD values and probability of similarity of sequences by chance.

| 2A: | | 2B: | |
|---|---|---|---|
| Atropinesterase compared with: | Score (SD units) | Probability of a score | Score (SD units) |
| trypsinogen (bovine) | -1.129 | $10^{-1}$ | 1.28 |
| α-lytic protease (Myxobacter 495) | 1.247 | $10^{-2}$ | 2.33 |
| subtilisin (B.amyloliquefaciens) | -0.128 | $10^{-3}$ | 3.09 |
| cephalosporinase (E.coli) | 0.982 | $10^{-4}$ | 3.72 |
| alkaline phosphatase (E.coli) | 0.352 | $10^{-6}$ | 4.75 |
| | | $10^{-8}$ | 5.61 |
| | | $10^{-10}$ | 6.36 |
| | | $10^{-15}$ | 7.94 |
| | | $10^{-20}$ | 9.26 |

family. The probabilities (expressed as SD values) found after comparison of atropinesterase with the serine-enzymes mentioned, are given in table VIII.2A. According to these values, atropinesterase could not be placed into either of the existing superfamilies. The highest SD value, obtained for the comparison with α-lytic protease, was 1.247. This SD value corresponds with a probability of similarity due to chance of $10^{-1}$. This probability-value was far too high to point to any relationship, as proteins with similar functions are not considered to be related unless the probability of similarity by chance is below $10^{-3}$ (SD > 3.1).

In agreement with the lack of evolutionary relationship between atropinesterase and the other serine-hydrolases is the position of the active serine in the polypeptide chain. In both identified superfamilies this residue occupies a position in the C-terminal third part of the protein, whereas in atropinesterase it is located at 40% from the N-terminus (see fig.VIII.2).

Summarizing: the methods used in the homology studies did not reveal the existence of any structural relationship between the newly sequenced serine-esterase atropinesterase and other known serine-enzymes. Even application of the program "RELATE", which could reveal a distant relationship between proteins of common ancestral origin, did not give any indication for such a relationship.

VIII.1.1.1 Experimental details

The computer programs "SEARCH" and "RELATE" (34) used in the homology studies are based on statistical data, i.e. the observed replacements of one amino acid by another in closely related sequences. These "accepted point mutations" seem to be imposed principally by natural selection and secondly by constraints of the genetic code. Some replacements by similar residues appear to have occurred frequently, e.g. Ile/Val and Lys/Arg, whereas 35 others of the 190 possible exchanges have never been observed. Obviously, the interchangeability has something to do with chemical similarity, with size, shape, electric charge or hydrophobicity. The observed exchanges (1572) were processed into a "scoring matrix" (38). For different evolutionary distances matrices were computed (distance was expressed in evolutionary intervals (PAMs); one PAM is defined as the interval in which one accepted point mutation per 100 residues has occurred). Each element of the "scoring matrix" for a given evolutionary distance gave a pair score: a figure which indicates that a pair of amino acids, one in either sequence, would be expected to occur in 2 related sequences (at the given evolutionary distance) more or less often than random chance would predict. In both programs "SEARCH" and "RELATE" the matrix of 250 PAMs was used, an evolutionary distance which corresponded with an observed difference of ca.80%. This scoring system has been proven most satisfactory for detecting distant relationships between sequences (119). The program "SEARCH" gave the pair scores for each pair of amino acids matched in the comparison. The sequence of 25 residues around the active serine of atropinesterase was compared with 246,737 segments of 25 residues derived from 1809 different proteins. By summing the pair scores, a search score was computed for each segment comparison. The maximum and minimum possible scores were 106 and -128, respectively, while the maximum and minimum scores observed from all comparisons were 45 and -59, respectively. The average score was -14.637, the standard deviation 11.058. The search produced no scores above the normal distribution which indicated that in all probability there were no sequences from relatives of atropinesterase in the database collection.

The matrix of 250 PAMs was also used in the program "RELATE". With this program the distribution of segment scores was obtained on comparison of all possible segments of 25 residues of atropinesterase with all segments of the same length from each of the reference proteins. A numerical property of the segment scores was calculated for the real sequences. The sequences were then randomly permuted and the same property was calculated for 100 randomized sequences. The score was calculated as the difference between the value determined for the real sequences and the average number determined from the many pairs of randomized sequences, divided by the standard deviation of the values from the randomized sequences. The numerical property of the segment comparison score was based on the average magnitude of a predetermined number of highest segment scores, i.e. in our case the number of scores to be expected if the sequences were related. In comparison of atropinesterase with longer proteins, this figure was: 248 (272-25+1), the number of 25 residues-segments in atropinesterase (in a comparison

with a related protein, at most all 248 segments can be matched, with a certain degree of correspondence). The segment comparison score, calculated for each pair of sequences, represented the probability of occurrence of that particular score by chance. The correlation between the SD values thus obtained and the probability of similarity by chance is given in table VIII.2B. To discriminate whether or not proteins descended from a common origin, Dayhoff imposes the following criteria: a newly determined sequence is placed in an existing superfamily if at least one probability of $< 10^{-6}$ (SD > 4.75) is obtained or if a probability of $< 10^{-3}$ (SD > 3.09) is obtained, when in advance it has been shown that the proteins share a similar function. The SD values found in our studies were far too low to indicate any similarity based on related sequences.

## VIII.2 PREDICTION OF PROTEIN STRUCTURE FROM THE AMINO ACID SEQUENCE

Although no sequence similarity appeared to exist between atropinesterase and other serine-hydrolases, there are distinct mechanistic and functional resemblances. The possibility was considered that evolutionary relationship with some enzymes still might exist, but that too many insertions, deletions and amino acid changes had occurred to defect the common ancestry. Additional evidence such as the nature of the three-dimensional structure could possibly permit the inference of relationship, as has been shown for the prothrombin-related enzymes α-lytic protease and elastase (67).

Knowledge of the tertiary structure and in particular the construction of the active center is also vital for understanding the biological function of the enzyme, such as the specificity for (−) atropine, and for insight in the reaction with compounds such as the organophosphorus inhibitors, since it is the unique three-dimensional structure of the protein which is essential for the recognition of possible substrates and which determines the way it functions. The primary structure of a protein in itself is not sufficient to understand the relation between structure and function, although -to a large extent- it determines the conformation i.e. the secondary and tertiary folding. Furthermore, knowledge of the primary structure is usually needed for the correct interpretation of data obtained in studies of the spatial shape of a protein. The detailed three-dimensional structure of a protein can only be established when suitable crystals can be grown.

Atropinesterase is one of the many proteins that defied all attempts made up to now to give appropriate crystals, so that other approaches had to be made to acquire structural information. (The tertiary structure is known of about 100 out of the more than 1800 proteins of which the sequence is known.) Some structural features can be determined by physico-chemical studies. Several aspects of the structure of the active center of atropinesterase were investigated by Van der Drift in our institute, in fluorescence, NMR and ESR studies on the enzyme labeled at the active center with appropriate reporter groups (42).

In principle prediction of the tertiary structure of a protein from its primary structure should be possible, as all the information necessary to acquire the native, biologically active conformation is dictated by the amino acid sequence (8). A possible approach to predicting tertiary structure from amino acid sequence was proposed by Sternberg et al. (125, 31). Starting with the primary sequence one could firstly predict the regular secondary structures, secondly pack the α-helices and β-strands into an approximate native fold and thirdly use simplified energy calculations to refine the fold into the native structure. The methods used so far to predict the secondary structure of proteins generally lead to structures with a moderate degree of accuracy (118). Predictions of the tertiary structure from the amino acid sequence, however, can not yet be performed in a reliable way because the prediction of the packing of α-helices (104, 30) and β-sheets (31, 102) still is a risky procedure. Therefore, even the construction of an approximated tertiary structure of relatively large proteins on the basis of the amino acid sequence is not feasible at present. Only when proteins can be grouped into families one can use the observed structure of one member of the family to predict the unknown structure of a protein with a related sequence in a reliable way, as has been shown for α-lytic protease (39).

Excellent reviews on protein structure and folding include those of Schulz and Schirmer (118), Rossman and Argos (110), Richardson (105) and Cantor and Schimmel (24).

## VIII-2.1 PREDICTION OF THE SECONDARY STRUCTURE

Many theories have been developed to predict the secondary structure of a protein from its amino acid sequence (for a review see Robson and Suzuki, 107). As discussed in detail by Burgess and Scheraga (22), the ability to make a good prediction of secondary structure is likely to remain a factor of crucial importance, whenever an attempt is made to predict the three-dimensional structure of a protein molecule. The secondary structures, such as α-helices and β-strands, are thought to be possible nucleation sites around which the remainder of the protein folds (see for example Robson and Pain, 106). Another important element of secondary structure is formed by the β-turns which consist of 4 residues and enable a polypeptide chain to turn by nearly $180^{o}$. Several investigators such as Chou and Fasman (28), Lewis et al. (76) and Venkatachalam (131) have examined the conformation of these turns. More than fifteen methods of predicting secondary structure are known to date (for some approaches, see Chou (26, 27), Robson and Suzuki (107), Lim (78), Ptitsyn (101), Nagano (87) and Garnier (53)). Most of the methods predict chain folding on the basis of statistical analysis of proteins of known sequence and conformation. These predictions assume that local sequence determines local structure. Usually only short-range interactions are considered, interactions between distant chain segments being ignored. The most important principle of these analyses is that the backbone conformation of an amino acid residue is determined, at least in part, by the type of its side chain. The empirical rules used in these studies are based on the distribution of single amino acids in their different combinations within the α-helical, β-structural regions and different bends of the protein chain. The method of Chou and Fasman (26, 27), as one of the statistical methods, has attracted much attention because it is relatively simple, can be applied without a computer, and has been relatively successful in its applications.

Some other methods do not incorporate statistical data, but are based on stereochemical criteria, like the method of Lim (78). This theory takes into account general "architectural" principles of packing of

polypeptide chains in globular proteins, such as the compactness of the molecule, the presence of tightly packed hydrophobic core (cores) and of a polar shell. These methods also take into account the interaction of proteins with water molecules; it considers the hydrophobicity and the size of the amino acid side chains and also incorporates, to some extent, long-range interactions. This a priori theory permits the identification of α-helical and β-structural regions in globular proteins without energy calculations.

Up to now the methods, sometimes applied in combination (116, 82, 45, 9), give predictions that are better than random for all structures, and are in many cases in good agreement with the experimentally determined structures. It has been suggested that it is the structural type of a protein that determines which predictive method will succeed (17). According to their secondary structure content, most proteins fall into one of 5 rather distinct classes (75). Proteins having only α-helices are designated as α-proteins, proteins with mainly β-strands are called β-proteins, α+β proteins have α-helix and β-strand segments that do not mix, whereas α/β proteins have mixed or approximately alternating segments of α-helix and β-strand; "coil" proteins contain almost no secondary structure.

Anyhow, none of the present methods or combinations thereof appears to have incorporated all the structural requirements for formation of α-helix, β-strand, β-turn or coil, as the percentage of correctly predicted structures (at least 80%; 31), has not markedly improved since 1974 although more protein structures are now available for verification. The fact that long-range interactions are not considered adequately, appears to be an important reason for the limitations in the accuracy that can be attained (125).

### VIII-2.1.1 PREDICTION OF THE SECONDARY STRUCTURE OF ATROPINESTERASE

For the prediction of the secondary structure of atropinesterase from the primary structure, we used the relatively complex method of Lim (78) which is based on stereochemical criteria, since fairly good correspondence was observed between the predicted secondary structures of

serine-enzymes such as trypsin, chymotrypsin and subtilisin and the experimental data.

VIII.2.1.1a The prediction theory of Lim. The physical theory of secondary structure of Lim has taken into account, as mentioned above, the main principles of the globular organization of protein chains. These principles are based on voluminous experimental data on three-dimensional structures (40) and intramolecular packing of water-soluble globular proteins (74, 79). In this method, the important structural features, such as the close-packing, the good hydrophobic burial and hydrophilic exposure of amino acids, impose structural requirements. Subsequently rules were proposed for α-helical and β-structural conformation.

By application of the proposed rules on the primary structure of 25 globular proteins, regular secondary structures were predicted for these molecules. These algorithms correctly identified α-helical and β-structural regions of all the examined proteins with an accuracy of about 80% and 85%, respectively and predicted correctly the state of about 70% of all the residues when simultaneously the requirements for all three types of secondary structure (irregular, α-helix, β-strand) were imposed. For example, after application of the proposed rules on the primary structure of the serine-enzymes subtilisin from Bacillus amyloliquefaciens (136), α-chymotrypsin (12) and trypsinogen (127), a good correspondence was observed between the predicted secondary structures and experimental data, as is shown graphically in fig. VIII.2. As is given in table VIII.3, on the average 89% of the residues of these enzymes were predicted correctly as α-helical or non-helical, 87% as β- or non-β-structural and 76% by three types of secondary structure. The predicted β-strands could be divided in internal type, surface type and semi-surface type β-strands according to the character of the residues in the β-strand.

Some shortcomings were observed in the prediction method, because the theory of Lim does not acknowledge α-helical and β-structural regions with a considerable deviation of their geometric parameters from the

standard values of the α-helix and β-structure. For instance, theoretical localization of α-helical and β-structural regions by this method can disagree with the experimental result, in regions of the chain where it enters the enzyme's active center. Furthermore, the method has the tendency to underestimate the number of residues in helical regions. Frequently this concerns helices with deviations in their geometric parameters from the standard values of the α-helix. With regard to β-structural regions, in some cases the prediction gives an overestimation compared to reality, in other cases the values are somewhat low.

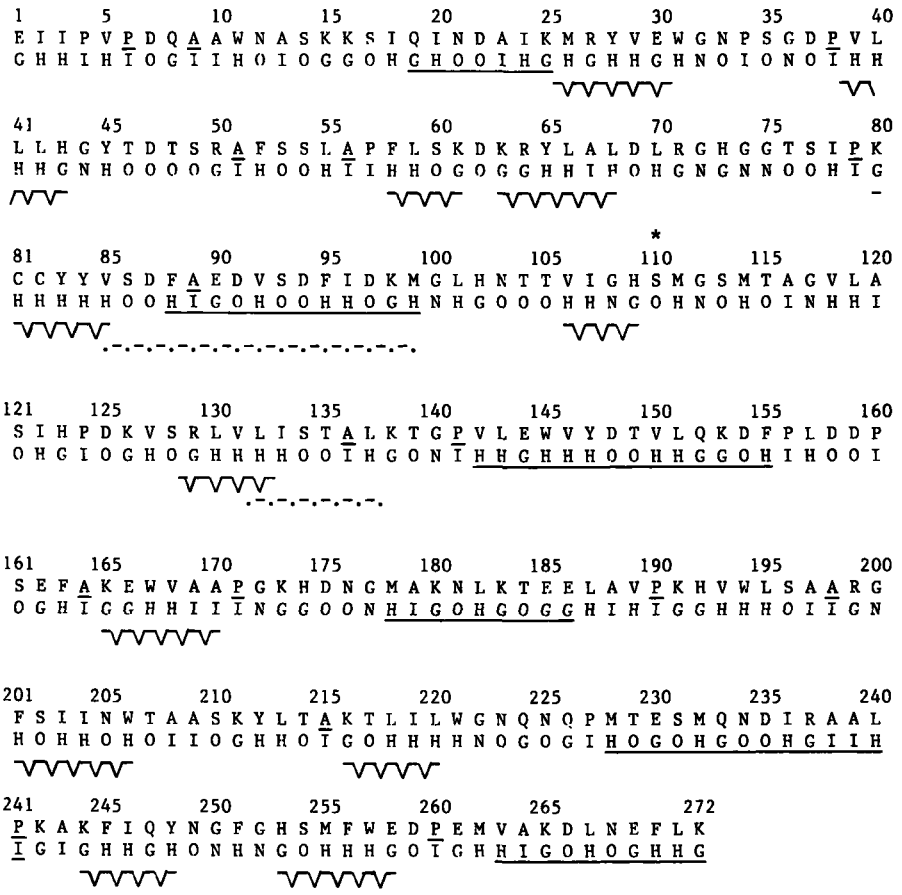
Despite these shortcomings when refinement of the results is at stake, in general the correspondence between prediction and experiment is very satisfactory. Therefore, the proposed rules were applied on the primary structure of atropinesterase.

VIII.2.1.1b The secondary structure of atropinesterase, predicted by the algorithm of Lim. Application of the rules on the primary structure of atropinesterase showed some regions of the chain to be α-helices, whereas other regions were designated as β-strands, as is shown in fig.VIII.1. There, the conformational states of the residues (α-helix and β-strand) are indicated by symbols. Atropinesterase is predicted to have 6 helical regions, involving the residues 19-25, 88-99, 142-155, 178-186, 228-241 and 263-272, while 12 β-strands were predicted, for the regions 26-31, 39-43, 58-61, 63-68, 80-85, 106-109, 129-133, 165-170, 201-206, 216-220, 244-248 and 253-258. Predicted β-strands: 95-98, 142-145, 150-153 and 269-272, which were completely overlapped by α-helical regions, and one β-strand, 261-265, which was partially overlapped by a helical region (263-272) are not mentioned, since, according to Lim, the helical conformation is energetically more favourable.

In a refinement of the method, certain Ala and Pro residues can be equated to large hydrophobic (H) residues; to which residues this applies depends on the sequence of which they form a part, but also certain long distance interactions have to be taken into consideration. The Pro and Ala residues in atropinesterase that obeyed these structural requirements were "promoted" to category H and the program

was run once more. The adaptation resulted in an extension of the α-helical region 88-99 to residues 85-99, which foreshortened the β-strand 80-85 to 80-84. Furthermore, a 7th α-helix was predicted for region 132-138, whereas the β-structural region 129-133 was obliterated.

Fig.VIII.1 Prediction of the secondary structure of atropinesterase




——— α-helix.
ϒϒ β-strand.
.-.-. α-helix, if the Ala and Pro residues underlined in the sequence are considered as large hydrophobic residues.
* active serine-residue

In the first row the amino acid sequence is presented in the one-letter code; in the second row each residue is designated with one of the symbols "H, I, G, O and N" for large hydrophobic, small hydrophobic, large hydrophilic, small hydrophilic and glycine, respectively (see VIII.2.1.1.1).

Table VIII.3 Helices and ß-strands predicted and experimentally determined in atropinesterase and other serine-enzymes.

| Enzyme | α-helical % of residues | | $Q_\alpha$ | ß-structural % of residues | | $Q_\beta$ | α- and ß-struct. % of residues | | $Q_{\alpha+\beta}$ | length of chain |
|---|---|---|---|---|---|---|---|---|---|---|
| | pred. | exp. | | pred. | exp. | | pred. | exp. | | |
| atropinesterase | 24 (28) | | | 24 (22) | | | 48 (50) | | | 272 |
| subtilisin | 18 | 31 | 82% | 13 | 10 | 93% | 31 | 41 | 76% | 275 |
| α-chymotrypsin | 5 | 7 | 94% | 19 | 29 | 88% | 24 | 36 | 82% | 245 |
| trypsinogen | 4 | 7 | 90% | 28 | 32 | 80% | 32 | 39 | 70% | 229 |

( ) : figures in parentheses, given for atropinesterase, are the percentages obtained when some Ala and Pro residues were equated to large hydrophobic residues.
pred: percentage of residues in the indicated state as predicted by the theory of Lim (78).
exp.: percentage of residues in the indicated state determined experimentally as calculated form the data published by Lim (78).
$Q_\alpha$, $Q_\beta$, $Q_{\alpha+\beta}$, percentages of residues the state of which is correctly predicted as helical or non-helical, ß- or non-ß-structural and by three types of secondary structure (irregular, α-helix and ß-strand), respectively (78).

According to the character of the residues predicted to occur in ß-strands, all strands in atropinesterase were characterized as internal type ß-structures, because they contained only hydrophobic amino acids but for one or two hydrophilic residues in the first two and (or) last two positions. For instance, Arg 27 and Glu 30 in strand 26-31 and His 43 were present near the termini of a ß-strand, in agreement with a requirement in the theory of Lim that side chains present in the hydrophilic shell of the protein should shield the local region of a hydrophobic core.

In the original approach, in total 66 residues (24% of all residues in the polypeptide chain) were predicted to be α-helical and 65 residues (24%) to be ß-structural. When selected Ala and Pro residues were equated to large hydrophobic residues 76 (28%) and 59 (22%) residues were predicted as α-helical and ß-structural, respectively.

VIII.2.1.1.1 Experimental details

The algorithm of Lim was applied on atropinesterase with the help of a computer program. The rules proposed by Lim were incorporated into the program in order to find regions of the chain compatible, in the α-helical or ß-structural conformation, respectively, with the structural requirements imposed by the theory of Lim (78). To simplify the search for regular secondary structures along the chain each amino acid was classified in a distinct group, according to the characteristics of the side group. This simplification is allowed since it has been shown that in general the peripheral parts of side groups of certain amino

acids are very similar in form and dimensions. The amino acids were classified and designated in the following way (see fig.VIII.1).
- Large hydrophobic residues, indicated with "H": Cys, ½(Cys-Cys), Val, Ile, Leu, Met, Phe, Tyr and Trp.
- Small hydrophobic residues, indicated with "I": Ala and Pro.
- Large hydrophilic residues, indicated with "G": Lys, Arg, Glu, Gln and His.
- Small hydrophilic residues, indicated with "O": Ser, Thr, Asp and Asn.
- Gly, heaving no side chain, has been indicated with "N".

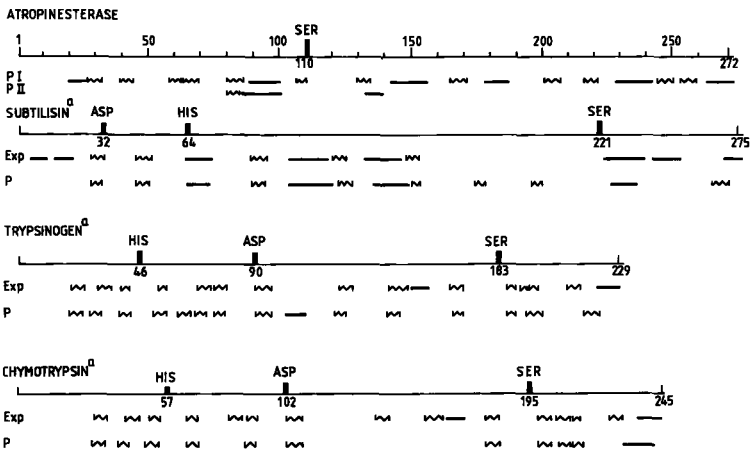
In fig.VIII.1, each residue along the polypeptide chain of atropinesterase is designated with one of these symbols.
To denote α-helices and β-strands along the polypeptide chain many interacting features were taken into account in the program. These features together with detailed restrictions, gave the information to predict the secondary structure. After matching of all the data to the rules proposed by Lim, the regions of the chain which in the helical conformation were compatible with the structural requirements, were assigned as α-helices and all the regions of the chain which satisfied the rules for β-structural conformation were assigned as β-strands. Some features of Lim were not incorporated in the original computer program, i.e. the equation of certain small hydrophobic residues (Ala and Pro) to large hydrophobic residues and the assignment in regions with competitive α- and β-structure. The first feature was incorporated as follows: in one computer run all Ala and Pro residues were considered as small hydrophobic residues (I), whereas in a second run only those Ala and Pro residues that obeyed rules A and B of Lim (78) were equated to large hydrophobic (H) residues. In both runs all possible α-helical and β-structural regions along the chain were predicted. Where competitive situations were encountered, or disturbance of form compactness, regions predicted as both α-helical and β-structural were assumed to be helical, as suggested by Lim.
The study of secondary structure prediction with the algorithm of Lim was kindly performed by J. v.d. Laan at the University of Groningen.

### VIII.2.2 COMPARISON OF THE PREDICTED SECONDARY STRUCTURE OF ATROPINESTERASE WITH THE PREDICTED AND EXPERIMENTALLY DETERMINED SECONDARY STRUCTURES OF SOME SERINE-ENZYMES

Comparison of the predicted secondary structure of atropinesterase with the predicted and experimentally obtained secondary structures of trypsinogen, α-chymotrypsin and subtilisin revealed firstly a striking difference in the total amount of α-helices (obtained by experiment and/or predicted) between subtilisin and atropinesterase on the one hand and the members of the trypsin-family on the other, as is shown in tabel VIII.3. The predicted amount of β-strand (24% of the total number of residues) in atropinesterase was higher than the predicted and observed amount of β-structure in subtilisin, but agreed rather

well with the amount of β-strand predicted for the trypsin-enzymes. In atropinesterase about 50% of the residues was predicted to be located in a regular secondary structure (α-helix and β-strand). This number was relatively high compared to the predicted and observed percentages of residues in a regular structure in the known serine-enzymes. According to the presence of a substantial number of both α-helices and β-strands, atropinesterase could be classified as an α/β-type protein (110). This is in agreement with the observation of Sternberg (125) that intracellular proteins usually fall into either the α- or the α/β-class. Subtilisin as well is an α/β-type protein. In contrast, trypsin and chymotrypsin, having mainly β-strands (fig.VIII.2), are considered as pure β-proteins. On basis of the overall secondary structures no homology was observed between atropinesterase and enzymes of the trypsin family. Despite the fact that atropinesterase and the members of the subtilisin-family seemed to belong to the same class of proteins (α/β-type), no homology was observed with respect to the distribution of the regular secondary structures along the

Fig.VIII.2 Predicted secondary structure of atropinesterase and comparison with predicted and experimental secondary structures of trypsin, chymotrypsin and subtilisin.




a) taken from Lim.
WW β-strands; —— α-helices.
Exp.: experimentally determined (upper row)
P : predicted by the theory of Lim (lower row)
In atropinesterase helices and β-strands were predicted considering all Ala and Pro as small hydrophobic residues (row I) and some Ala and Pro-residues as large hydrophobic residues (row II), respectively.

polypeptide chain. Also when the active serine-residues of the enzymes were matched no homology in secondary structure could be observed.

### VIII.2.3 PREDICTION OF THE TERTIARY STRUCTURE OF ATROPINESTERASE

Because of the complete absence of relationship between atropinesterase and other proteins according to their primary and secondary structures, together with the lack of reliable programs, a prediction of the tertiary structure has not been attempted.

## VIII.3 CONCLUSIONS

No homology could be established with regard to the primary structure between atropinesterase and any of the other serine-enzymes. Structural relationship with respect to the content and kind of secondary structure, or with regard to the distribution along the polypeptide chain could not be observed either. Consequently, the conclusion appears inevitable that atropinesterase of Pseudomonas putida, strain PMBL-1, is the first structured member of a new superfamily of serine-hydrolases. Evidently, the independent evolutionary development of an esterolytic enzyme that owes its enzymatic activity to an activated OH of one particular serine residue has occurred at least three times. It remains to be proven whether this convergential evolution of different types of serine-hydrolases has resulted in a virtually absolute convergence with respect to the construction attained for the activation of the serine residue, or that other constructions can occur. Also the question whether or not the serine-proteases and serine-esterases show mechanistic and structural differences remains to be solved.

ACKNOWLEGMENTS

Many thanks to W.C.M. Zwennis for his valuable contribution to this study by performing the sequence analyses of the cyanogen bromide peptides of the protein and the stimulating discussions. Furthermore, I want to express my gratitude to Dr. F. Berends and Dr. R.A. Oosterbaan for their enthusiastic interest in my work and Dr. F. Berends for critical reading of the manuscript. F.J.A. Kouwenberg is gratefully acknowledged for isolation of the enzyme and thanks are due to all persons who gave technical assistance.

LITERATURE


1 Adelstein, R.S., and Kuehl, W.M., (1970) Biochemistry 9, 1355-1364.
2 Alden, R.A., Wright, C.S., and Kraut, J., (1970) Phil.Trans.Royal Soc.London B257, 119-224.
3 Ambler, R.P., (1963) Biochem.J. 89, 349-378.
4 Ambler, R.P., (1972) Methods in Enzymol.(Hirs, C.H.W., Ed.) Vol.XXVb, 143-154, 262-272 New York: Academic Press.
5 Ambler, R.P., and Brown, L.H., (1967) Biochem.J. 104, 784-825.
6 Ambler, R.P., and Meadway, R.J., (1968) Biochem.J. 108, 893-895.
7 Ambler, R.P., and Wynn, M., (1973) Biochem.J. 131, 485-498.
8 Anfinsen, C.B., Science (1973) 181, 223-230.
9 Argos, P., Schwarz, J., and Schwarz, J., (1976) Biochim.Biophys. Acta 439, 261-273.
10 Bennett, J.C., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 330-339 New York: Academic Press.
11 Berends, F., (1964) Thesis, Leiden p. 12.
12 Birktoft, J.J., and Blow, D.M., (1972) J.Miol.Biol. 68, 187-240.
13 Blömback, B., Blömback, M., Hessel, B., and Wanaga, S., (1967) Nature (Lond.) 215, 1445-1448.
14 Blow, D.M., Birktoft, J.J., and Hartley, B.S., (1969) Nature (Lond.) 221, 337-340.
15 Bode, W., and Schwager, P., (1975) J.Mol.Biol. 98, 693-717.
16 Boigne, J.M., Boigne, N., and Rosa, J., (1970) J.Chromat. 47, 238-246.
17 Bradshaw, R.A., Cancedda, F., Ericsson, L.H., Neumann, P.A., Piccoli, S.P., Schlesinger, M.J., Shriefer, K., and Walsh, K.A., (1981) Proc.Nat.Acad.Sci., U.S.A. 78, 3473-3477.
18 Braunitzer, G., Schrank, B., Ruhfus, A., Petersen, S., and Petersen, U., (1971) Hoppe-Seyler's Z.Physiol.Chem. 352, 1730-1732.
19 Brayer, G.D., Delbaere, L.T.J., and James M.N.G., (1979) J.Mol. Biol. 131, 743-775.
20 Brenner, M., Niederweiser, A., and Pataki, G., (1961) Experientia 17, 145-192.
21 Bridgen, J., (1977) Science Tools, The LKB Instrument Journal 24(1), 1-6.
22 Burgess, W.A., and Scheraga, H.A., (1975) Proc.Nat.Acad.Sci., U.S.A., 72, 1221-1225.
23 Butler, P.J.G., Harris, J.I., Hartley, B.S., and Leberman, R., (1969) Biochem.J. 112, 679-689.
24 Cantor, C.R., and Schimmel, P.R., (1980) Biophysical Chemistry Part I: San Francisco: Freeman, 365 pp.
25 Catravas, G.N., (1964) Anal.Chem. 36, 1146-1148.
26 Chou, P.Y., and Fasman, G.D., (1974) Biochemistry 13, 211-222.
27 Chou, P.Y., and Fasman, G.D., (1974) Biochemistry 13, 222-245.
28 Chou, P.Y., and Fasman, G.D., (1977) J.Mol.Biol. 115, 135-175.
29 Cleland, W.W., (1964) Biochem. 3, 480-482.
30 Cohen, F.E., Richmond, T.J., and Richards, F.M., (1979) J.Mol. Biol. 132, 275-288.

31 Cohen, F.E., Sternberg, M.J.E., and Taylor, W.R., (1980) Nature 285, 378-382.
32 Cohen Solal, M., and Bernard, J.L., (1973) J.Chromat. 80, 140-143.
33 Crestfield, A.M., Moore, S., and Stein, W.H., (1963) J.Biol.Chem. 238, 622-627.
34 Dayhoff, M.O., (1978) Atlas of Protein Sequence and Structure (Dayhoff, M.O., Ed.) Vol.5 suppl.3, 1-8, Nat.Biomed.Res.Foundation, Washington D.C., U.S.A.
35 Dayhoff, M.O., Barker, W.C., and Hardman, J.K., (1972) Atlas of Protein Sequence and Structure (Dayhoff, M.O., Ed.) Vol.5, 53-66 Nat.Biomed. Res.Foundation, Washington D.C., U.S.A.
36 Dayhoff, M.O., Barker, W.C., Hunt, L.T., and Schwartz, R.M., see ref. 34, 9-24.
37 Dayhoff, M.O., Hunt, L.T., Barker, W.C., Orcutt, B.C., Yeh, L.S., Chen, H.R., George, D.G., Blomquist, M.C., Fredrickson, J., and Johnson, G.C., (1982) Protein sequence database from the Atlas of Protein Sequence and Structure, version 4, Nat.Biomed.Res.Foundation, Washington D.C., U.S.A.
38 Dayhoff, M.O., Schwartz, R.M., and Orcutt, B.C., (1978) see ref. 34, 345-352.
39 Delbaere, L.T.J., Brayer, G.D., and James, M.N.G., (1979) Nature 279, 165-168.
40 Dickerson, R.E., and Geis, I., The structure and action of proteins, Harper and Row, New York, Evanston, London.
41 Dreyer, W.J., and Bynum, E., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 37 New York: Academic Press.
42 Drift, A.C.M. van der, Thesis in preparation, Utrecht.
43 Dubois, M., Gilles, K.A., Hamilton, J.R., Rebers, P.A., and Smith, F., (1956) Anal.Chem. 28, 350-356.
44 Dunker, A.K., and Ruekert, R.R., (1969) J.Biol.Chem. 244, 5074-5080.
45 Dzionara, M., Robinson, S.M.L., and Wittmann-Liebold, B., (1977) Hoppe-Seyler's Z.Physiol.Chem. 358, 1003-1019.
46 Easly, C.W., Zegers, B.J.M., and Vijlder, M. de (1969) Biochim. Biophys. Acta 175, 211-213.
47 Edman, P., (1950) Acta Chem.Scand. 4, 283-293.
48 Edman, P., (1977) Carlsberg Res.Commun. 42, 1-9.
49 Edman, P., and Begg, G., (1967) Eur.J.Biochem. 1, 80-91.
50 Frank, G., (1979) Hoppe-Zeyler's Z.Physiol.Chem. 360, 997-999.
51 Freedman, M.H., Grossberg, A.L., and Pressman, D., (1968) Biochemistry 7, 1941-1950.
52 Fricke, U., (1975) Anal.Biochem. 63, 555-558.
53 Garnier, J., Osguthorpe, D.J., and Robson, B., (1978) J.Mol.Biol. 120, 97-120.
54 George, D.G., Orcutt, B.C., Dayhoff, M.O., and Barker, W.C., (1982) NBR Report 820502-08710 Nat.Biomed.Res.Foundation, Washington D.C., U.S.A.
55 Giglio, J.R., (1977) Anal.Biochem. 82, 262-264.
56 Gray, W.R., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 139 New York: Academic Press.

57 Gray, W.R., (1972) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XXVb, 333-344 New York: Academic Press .
58 Gros, C., and Labouesse, B., (1969) Eur.J.Biochem. 7, 463.
59 Gross, E., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 238-242 New York: Academic Press.
60 Hartley, B.S., (1970) Biochem.J. 119, 805-822.
61 Hartley, B.S., (1970) Phil.Trans.Royal Soc. London B257, 77-87.
62 Hartley, B.S., Brown, J.R., Kauffman, D.L., and Smillie, L.B., (1965) Nature 207, 1157-1159.
63 Hill, R.L., and Delaney, R., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 339-351 New York: Academic Press.
64 Hirs, C.H.W., (Ed.) (1967) Methods in Enzymol. Vol.XI, 325-329 New York: Academic Press (idem 32b).
65 Hirs, C.H.W., Timasheff, S.N., (Eds.) (1977) Methods Enzymol. Vol.47, 668 pp.
66 Houmard, J., Drapeau, G.R., (1972) Proc.Nat.Acad.Sci. U.S.A. 69, 3506-3509.
67 James, M.N.G., Delbaere, L.T.J., and Brayer, G.D., (1978) Can.J. Biochem. 56, 396-402.
68 Jaurin, B., and Grundstrom, T., (1981)Proc.Nat.Acad.Sci. U.S.A. 78, 4897-4901.
69 Kauffman, D.L., (1965) J.Miol.Biol. 12, 929-932.
70 Kawasaki, I., and Itano, H.A., (1972) Anal.Biol.Chem. 48, 546-556.
71 Kikuchi, Y., Yoda, K., Yamasaki, M., and Tamura, G., (1981) Nucl.Acids Res. 9, 5671-5678.
72 Klapper, D.G., Wilde III, C.E., and Donald Capra, J., (1978) Anal.Biochem. 85, 126.
73 Konigsberg, W.H., and Steinman, H.M., (1977) The Proteins, Vol.3, Neurath, H., Hill, R.L., (Eds.) 1-178. New York: Academic. 663 pp. 3rd ed. (1976) J.Mol.Biol. 106, 983-994.
74 Lee, B., and Richards, F.M., (1971) J.Mol.Biol. 55, 379-400.
75 Levitt, M., and Chotia, C., (1976) Nature 261, 552-558.
76 Lewis, P.N., Momany, F.A., and Scheraga, H.A., (1971) Proc.Nat. Acad.Sci. U.S.A. 68, 2293-2297.
77 Light, A., (1972) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol. XXVb, 253-262 New York: Academic Press.
78 Lim, V.I., (1974) J.Mol.Biol. 88, 857-872, 873-894.
79 Lim, V.I., and Ptitsyn, O.B., (1972) Biofyzika (U.S.S.R.) 17, 21-33.
80 Markland, F.S., and Smith, E.L., (1967) J.Biol.Chem. 242, 5198-5211.
81 Matsubara, H., and Sasaki, R.M., (1969) Biochem.Biophys.Res.Comm. 35, 175-181.
82 Matthews, B.W., (1975) Biochim.Biophys.Acta 405, 442-451.
83 Maurer, H.R., Disk-Elektroforese, Theorie und Praxis der Diskontinuierlichen Polyacrylamide Gel Elektrophoresis.
84 Michl, H., (1951) Monatshefte für Chem. 82, 489-493.
85 Mikes, O., Holeysovsky, V., Tomasek, V., and Sorm, F., (1966) Biochem.Biophys.Res.Commun. 24, 316-352.

86 Mitchell, W.M., and Harrington, W.F., (1968) J.Biol.Chem. 243, 4683-4692.
87 Nagano, K., (1977) J.Mol.Biol. 109, 251-274.
88 Nakai, N., Lai, C.Y., and Horecker, B.L., (1974) Anal.Biochem. 58, 563-570.
89 Needleman, S.B., Ed. (1975) Protein Sequence Determination. New York: Springer. 393 pp.
90 Needleman, S.B., Ed. (1977) Advanced Methods in Protein Sequence Determination. New York: Springer. 189 pp.
91 Niall, H.D., (1973) Methods in Enzymol. XXVII, 942-1010.
92 Niall, H.D., (1977) see Ref.57, 179-238.
93 Offord, R.E., (1966) Nature (Lond.) 211, 591-593.
94 Olaitan, S.A., de Lange, R.J., and Smith, E.L., (1968) J.Biol. Chem. 243, 5296-5301.
95 Olson, M.O.J., Nagabhushan, N., Dzwiniel, M., Smillie, L.B., and Withaker, D.R., (1970) Nature 228, 438-442.
96 Oosterbaan, R.A., (1969) Pharmaceutical "Enzymes and their assay", Universitaire Pers, 10-21.
97 Oosterbaan, R.A., and Berends, F., (1971) Proceedings of the Royal Dutch Academy of Sciences, Series C, 74, 158-166.
98 Perham, R.N., Ed., (1975) Instrumentation in Amino Acid Sequence Analysis. London: Academic. 197 pp.
99 Perham, R.N., (1976) FEBS Letters Supplement 62, E20-E29.
100 Peterson, J.D., Nehrlich, S., Oyer, P.E., and Steiner, D.F., (1972) J.Biol.Chem. 247, 4866-4871.
101 Ptitsyn, O.B., and Finkelstein, A.V., (1970) Biofyzika (USSR) 15, 757-767.
102 Ptitsyn, O.B., Finkelstein, A.V., and Falk (Bendzko) P., (1979) FEBS Lett. 101, 1-5.
103 Ramachandran, L.K., and Witkop, B., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 283-299 New York: Academic Press.
104 Richmond, T.J., and Richards, F.M., (1978) J.Mol.Biol. 119, 537-555.
105 Richardson, J.S., Adv.Protein Chem. in press.
106 Robson, B., and Pain, R.H., (1976) Biochem.J. 155, 331-344.
107 Robson, B., and Suzuki, E., (1976) J.Mol.Biol. 107, 327-356.
108 Rörsch, A., Berends, F., Bartlema, H.C., and Stevens, W.F., (1971) Proceedings of the Royal Dutch Academy of Sciences, Series C, 74, 132-147.
109 Roseau, G., and Pantel, P., (1969) J.Chromat. 44, 392-395.
110 Rossmann, M.G., and Argos, P., (1981) Ann.Rev.Biochem. 50, 497-532.
111 Sanger, F., Proc.Chem.Soc. (1963) vol. no., 76-83.
112 Scheffer, A.J., (1973) Thesis, Groningen.
113 Schlesinger, D., (1979) Anal.Biochem. 95, 494.
114 Schroeder, W.A., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 445-461 New York: Academic Press.
115 Schroeder, W.A., Jones, R.T., Cormick, J., and McCall, K., (1962) Analyt.Chem. 34, 1570-1575.

116 Schulz, G.E., Barry, C.D., Friedman, C., Chou, P.Y., Fasman, G.D., Finkelstein, A.V., Lim, V.I., Ptitsyn, O.B., Kabat, R.A., Wu, T.T., Levitt, M., Robson, B., and Nagano, K., (1974) Nature 250, 140-142.
117 Schulz, G.E., (1977) Angew.Chemie (Intl. Edn.) 16, 23-32.
118 Schulz, G.E., and Schirmer, R.H., Principles of Protein Structure (1979) (Springer-Verlag, New York).
119 Schwartz, R.M., and Dayhoff, M.O., (1978) see ref.34, 353-358.
120 Smithies, O., Gibson, D., Fanning, E.M., Goodfliesh, R.M., Gilman, J.G., and Ballantyne, D.L., (1971) Biochemistry 10, 4912-4921.
121 Smyth, D.G., (1967) Methods in Enzymol. (Hirs, C.H.W., Ed.) Vol.XI, 216-222 New York: Academic Press.
122 Spackman, D.H., Stein, W.H., and Moore, S., (1958) Anal.Chem. 30, 1190-1206.
123 Spackman, D.H., (1963) Fed.Proc.Fed.Amer.Soc.Exp.Biol. 22, 244.
124 Stein, S., Böhlen, P., Stone, J., Dairman, W., and Udenfriend, S., (1973) Arch.Biochem.Biophys. 155, 202-212.
125 Sternberg, M.J.E., and Thornton, J.M., (1978) Nature 271, 15-20.
126 Stevens, W.F., (1969) Thesis, Leiden.
127 Stroud, R.M., Kay, L.M., and Dickerson, R.E., (1972) Cold Spring Harbor Symp.Quant.Biol. 36, 125-140.
128 Tarr, G.E., Beecher, J.F., Bell, M., and McKean, D.J., (1978) Anal.Biochem. 84, 622-627.
129 Titani, K., Ericsson, L.H., Neurath, H., and Walsh, K.A., (1975) Biochemistry 14, 1358-1366.
130 Udenfriend, S., Stein, S., Böhlen, P., Dairman, W., Leimgruber, W., and Weigele, M., (1972) Science 178, 871-872.
131 Venkatachalam, C.M., (1968) Biopolymers 6, 1425-1436.
132 Walsh, K.A., Ericsson, L.H., Parmelee, D.C., and Titani, K., (1981) Ann.Rev.Biochem. 50, 261-284.
133 Walsh, K.A., Ericsson, L.H., and Titani, K., (1978) Versatility of Proteins, Li, C.H., (Ed.) 39-58. New York: Academic.
134 Weber, K., and Osborn, M., (1969) J.Biol.Chem. 244, 4406-4412.
135 Woods, K.R., and Wang, K.T., (1967) Biochim.Biophys.Acta 133, 369-370.
136 Wright, C.S., Alden, R.A., and Kraut, J., (1969) Nature (Lond.) 221, 235-242.
137 Yamada, S., and Itano, H.A., (1966) Biochim.Biophys.Acta 130, 538-540.
138 Young, C.L., Barker, W.C., Tomaselli, C.M., and Dayhoff, M.O., (1978) ref. 76, 73-93.
139 Zimmerman, C.L., Appella, E., and Pisano, J.J., (1977) Anal.Biochem. 77, 569.

SAMENVATTING

Acetylcholinesterase is een estersplitsend enzym dat een essentiële functie heeft in het zenuwstelsel; bij remming van dit esterase wordt de overdracht van zenuwprikkels ontregeld of geheel geblokkeerd. Acetylcholinesterase is zeer gevoelig voor bepaalde organische fosforverbindingen die daarom kunnen worden toegepast als insecticiden of chemische strijdmiddelen. Deze "organofosfaten" reageren specifiek met de z.g. aktieve serine van het enzym, dat daardoor zijn aktiviteit verliest. Deze reaktie is typerend voor een groep hydrolytische enzymen met een verwant katalytisch mechanisme, de z.g. serine-esterasen. Om meer inzicht te krijgen in het werkingsmechanisme van deze enzymen en in hun reaktie met organofosfaten is kennis omtrent de struktuur van belang. Van verschillende serine-enzymen, zoals subtilisine en de aan elkaar verwante enzymen trypsine en chymotrypsine, is de struktuur al opgehelderd. Hoewel ze ook estersplitsing katalyseren, zijn dit in wezen echter eiwitsplitsende enzymen (proteasen), in tegenstelling tot acetylcholinesterase dat een zuiver esterase is. Dit verschil zou kunnen berusten op verschillen in struktuur van het katalytisch systeem. Acetylcholinesterase is slecht toegankelijk voor biochemisch en biofysisch onderzoek, o.a. door de grootte van het molecuul. Dit geldt ook voor andere bekende serine-esterasen. Voor het onderzoek naar de struktuur van een esterase werd daarom gebruik gemaakt van het model-enzym atropinesterase. Dit echte serine-esterase, dat de esterbinding in het alkaloid atropine verbreekt, wordt geproduceerd door Pseudomonas putida bacteriën, die geïsoleerd zijn uit grond die zich bevond tussen de wortels van de Atropa belladonna. Atropinesterase is een dimeer. De identieke monomeren bestaan uit één polypeptide keten van 272 aminozuur-residuen, met één aktieve serine (hoofdstuk IV). De primaire struktuur van het eiwit, dat is de volgorde van de aminozuren, werd vastgesteld ten einde te kunnen bepalen of dit serine-esterase verwant is aan bekende serine-proteasen. Hiertoe moest het eiwit op specifieke wijze in een beperkt aantal fragmenten worden gesplitst; hierbij werd gebruik gemaakt van een enzymatische afbraakmethode met trypsine en een chemische afbraakmethode met cyanogeenbromide. Allereerst werden peptiden geïsoleerd die de aktieve

serine omvatten. Vergelijking van de aminozuurvolgorde van het met cyanogeenbromide verkregen peptide (met 12 residuen), met de volgorde rondom de aktieve serine in trypsine en subtilisine was onbevredigend, omdat in dit fragment de aktieve serine gelegen was op positie 11, waardoor vergelijking slechts aan één zijde van dit residu mogelijk was. Afbraak van gemaleyleerd atropinesterase met trypsine gaf een fragment met 58 residuen met de serine op positie 39. Dit fragment vertoonde geen homologie met overeenkomstige gebieden van bekende serine-enzymen. De isolatie, zuivering en sequentiebepaling van deze peptiden zijn beschreven in hoofdstuk V.

Om na te gaan of er sprake was van enige evolutionaire verwantschap werd vervolgens de aminozuurvolgorde van het gehele eiwit vastgesteld. In hoofdstuk VI is de isolatie beschreven van de 8 peptiden, verkregen na afbraak van het gemaleyleerde eiwit met trypsine, evenals de isolatie van 6 peptiden gevormd bij afbraak met cyanogeenbromide en van 2 met pepsine verkregen fragmenten. Hoofdstuk VII beschrijft hoe de aminozuurvolgorde van het gehele eiwit kon worden vastgesteld, nadat de aminozuurvolgorde van alle in hoofdstuk VI beschreven peptiden was opgehelderd. Hierbij werd gebruik gemaakt van conventionele sequentieanalyse technieken, waar mogelijk aangevuld met meer geavanceerde technieken. De verkregen resultaten leverden alle informatie voor het vaststellen van de volgorde van de 272 aminozuren, zoals weergegeven in tabel VII-VI. Atropinesterase is het eerste echte serine-esterase waarvan de primaire struktuur is opgehelderd.

Het onderzoek naar mogelijke gelijkenis wat betreft de primaire struktuur tussen atropinesterase en bekende serine-proteasen (trypsine en subtilisine), volgens een computerprogramma van Dayhoff, toonde geen verwantschap. Ook een voorspelling omtrent de secundaire struktuur, de ruimtelijke bouw van gedeelten van de eiwitketen, liet geen overeenkomst zien (hoofdstuk VIII). Evolutionair gezien lijkt atropinesterase derhalve afkomstig te zijn van een ander "oerenzym" dan de reeds bekende serine-enzymen, de subtilisines en de aan elkaar verwante trypsines en chymotrypsines. Atropinesterase vormt hiermee het eerste bekende lid van een derde superfamilie van serine-hydrolasen.

CURRICULUM VITAE

De schrijfster van dit proefschrift werd op 18 oktober 1943 geboren te Voorburg. Op verzoek van de Faculteit der Wiskunde en Natuurwetenschappen volgt hier een overzicht van het verloop van haar studie.
In 1962 werd het eindexamen gymnasium β behaald aan het St.-Maartenslyceum te Voorburg. Na een korte onderbreking begon zij in 1963 met de opleiding voor leerling-analiste te Scheveningen, waarvoor in 1964 het diploma werd behaald. In hetzelfde jaar werd begonnen met de studie in de scheikunde aan de Rijksuniversiteit te Leiden. Het kandidaatsexamen, letter g, werd in november 1967 afgelegd. Het doctoraalexamen met als hoofdvak organische chemie en als bijvak pathologische chemie werd cum laude afgelegd in juni 1970; de studie hiervoor stond onder leiding van de hoogleraren Dr. E. Havinga en Dr. A.A.H. Kassenaar.
Vanaf december 1970 is zij als wetenschappelijk medewerkster verbonden aan het Medisch Biologisch Laboratorium van de Hoofdgroep Gezondheidsonderzoek (HGO) TNO te Rijswijk, waar onder leiding van Dr. F. Berends en Dr. R.A. Oosterbaan het hier beschreven onderzoek werd verricht.

NAWOORD

Gaarne wil ik iedereen, die mij op enigerlei wijze in de gelegenheid heeft gesteld dit proefschrift tot stand te brengen en hiermee mijn opleiding te voltooien van harte bedanken. Niet alleen jullie daadwerkelijke hulp heeft bijgedragen tot dit resultaat, ook jullie morele steun, welke ik in de vele stadia van dit onderzoek heb ondervonden en welke ik als zeer waardevol heb ervaren. Bedankt allemaal voor de gezamenlijke inspanning waarvan hier het bewijs.

Wellicht kan de ten toon gespreide coöperatie een voorbeeld zijn voor de wenselijke samenwerking tussen onderzoekers op het gebied van de eiwitsequentie, de DNA-sequentie en de röntgenkristallografie om een antwoord te vinden op de vraag welke magische rol de 20 aminozuren spelen in levende organismen.

ENZYMES ARE PROTEINS,
THINGS OF BEAUTY
AND A JOY FOR EVER.