

University of Groningen

The structure of papain

Koekoek, Roelof

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1969

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Koekoek, R. (1969). *The structure of papain: An x-ray diffraction study at 2.8. Å resolution.* s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

SUMMARY

The work described in this thesis is a part of the study of the proteolytic enzyme papain (EC 3.4.4.10) by means of X-ray diffraction. The first stages of this study, described by Drenth *et al* [53, 58] and by Jansonius [1], resulted in the calculation of an electron density map at a resolution of 4.5 Å. From this map the papain molecules could be distinguished and their shape could be determined as being more or less spheroidal with dimensions 36 x 48 x 36 Å. However, owing to the large number of intersections, the map did not show the course of the polypeptide chain.

In the present work an electron density map was calculated to a resolution of 2.8 Å. This map completely revealed the tertiary structure of the enzyme. In addition, the chemically determined amino-acid sequence [25] was corrected by transposing a string of 40 amino-acids *in toto* from the second part of the sequence to the first part. Thirteen residues had to be inserted in the position of non-overlap in the chemically determined sequence. With an extra valine residue, found at a different position, the total number of amino-acid residues thus became 212. The nature of the 13 unknown residues has been established recently by Lowe [80]. The molecular weight of papain could now be calculated to be 23,350.

The polypeptide chain is folded in such a way that two distinct parts, which are divided by a cleft, can be distinguished. In this cleft the active site of the enzyme, with a cysteine and a histidine residue as the essential constituents, is located. These residues are found opposite each other on either side of the cleft. The distance between their side chains is about 4 Å.

In Chapter I a short review of the numerous chemical studies on papain is given. Papain is a member of the class of enzymes which need a free sulphydryl group for activity. The enzyme acts as a catalyst in the hydrolysis of certain peptides, amides and esters. Its specificity is not very well defined since it can hydrolyze peptide bonds formed by almost

all of the amino-acids. A slight preference for the splitting of the bonds of the general type Lys-X or Arg-X was assumed to exist [12], but recently Schechter and Berger [13] found an even stronger preference for splitting the X-Y bond in peptides in which a phenylalanine residue precedes the residues X and Y, the nature of X and Y being unimportant (-Phe-X[↓]Y-).

In papain, prepared according to the method by Kimmel and Smith [9], the SH-group is barely present as such. One part of the molecules has reacted with cysteine to yield a mixed disulphide [26, 62]. Of the other part, the SH-group is irreversibly blocked presumably by oxidation. The enzyme can be activated by converting the mixed disulphide into the free SH-group, e.g. by an excess of cysteine. In addition, since papain is inhibited strongly by traces of metals, a metal chelating agent such as EDTA should be added.

In Chapter II the various crystal modifications of papain are described (Table IV). The orthorhombic form C, grown from 70% methanol, was used for the diffraction study. The enzyme is inactive in this solution. However, it can be shown by various methods [51, 118, 136] that the conformation of papain in the crystallization medium is not significantly different from the native conformation.

The phase angles of the reflections were obtained by the isomorphous replacement technique. Therefore a series of heavy-atom containing derivatives of papain was prepared. By measuring the changes in intensities of corresponding reflections the phase angles could be derived. Five derivatives, containing PCMB, PCMS, PCMA, (Pap-S)-HgCl and (Pap-S)-HgCl + HgCl₂ respectively, were used (Table V). In addition, the anomalous scattering by the mercury atoms in the derivatives was successfully used in the determination of the phase angles.

In Chapter II is also described how the intensities of the diffracted X-ray beams were measured. In the low resolution work this was done by taking precession photographs. In the present work an automatic diffractometer was used. Some characteristic features of both the Arndt and Phillips linear diffractometer [126] and the triple-channel version of this instrument [127, 128] are described. In addition, the semi-

empirical absorption corrections, which were applied, and the scaling of intensities of different crystals are discussed.

In Chapter III the evaluation of the phase angles from the intensity data is described. The mean "figure of merit" [151] finally obtained was 0.82 for 9525 reflections. This is quite satisfactory when compared to other protein structure determinations. The influence of anomalous scattering differences on the phase angles was found to be important and, therefore, these data were included in the phase calculation.

In Chapter IV the interpretation of the "best" [150] electron density map is described. The side chains are all clearly visible as regions of high density pointing sideways away from the continuous density of the main chain. The resolution proved to be sufficient to recognize side chains such as those of tryptophan, tyrosine, cysteine, isoleucine etc., which have characteristic features. In the main chain almost all carbonyl groups can be identified as the regions of relatively high density.

Five pieces of right handed α -helix are observed in the papain molecule. Together they comprise about 60 residues or 28% of the total number. This is in good agreement with the α -helical content predicted by the most recent ORD measurements [201]. A number of about 32 residues occur in a so-called twisted pleated sheet structure (fig. 24). Only about 10 of these residues form a regular antiparallel pleated sheet or β -structure [177].

A short description of the active site components is given in section 4.3. Also included are the recently suggested [37,185] mechanisms of papain catalysis.

Finally the results of some difference Fourier maps of modified papain compared to the native protein are discussed. First, the derivative of papain, prepared by Wolthers [189], in which the SH-group has reacted with tosyl-lysyl-chloro ketone (TLCK), is described. Secondly, certain amino groups of the protein were labelled by reaction with the Edman-reagent phenylisothiocyanate as well as with the iodinated form of this same compound: *p*-iodophenylisothiocyanate. The positions of the sulphur and iodine atoms could be determined. It turned out that 3 lysines and the α -NH₂ group of the protein had reacted. In the third place, tyrosine

residues were iodinated by reaction with KI_3 . At least 4 out of 7 peaks observed in the difference Fourier maps of two projections can be correlated with tyrosine residues.

These results provide an additional confirmation of the correctness of the amino-acid sequence as deduced from the electron density map (fig. 21).

SAMENVATTING

Het in dit proefschrift beschreven onderzoek maakt deel uit van de bepaling van de structuur van het eiwitsplitsend enzyme papaine (EC 3.4.4.10), door middel van röntgen-diffractie. De eerste stadia van de structuurbepaling zijn beschreven door Drenth c.s. [53, 58] en door Jansonius [1] en resulteerden in de berekening van een electronendichtheidsverdeling met een oplossend vermogen van $4,5 \text{ \AA}$. In deze dichtheidsverdeling konden de moleculen worden onderscheiden en hun ruwe vorm kon worden bepaald als ellipsoïden met afmetingen $36 \times 48 \times 36 \text{ \AA}$. De loop van de polypeptideketen kon echter niet worden vastgesteld, als gevolg van een groot aantal "valse" contacten.

In dit proefschrift wordt beschreven hoe uit een electronendichtheidsverdeling met een oplossend vermogen van $2,8 \text{ \AA}$ de tertiaire structure van het enzyme volledig kon worden bepaald. Bovendien werd de chemisch bepaalde aminozuurvolgorde [25] gecorrigeerd door het verplaatsen van een uit 40 aminozuren bestaand peptide van de tweede helft van de volgorde naar de eerste helft. Dertien aminozuren moesten aan de volgorde worden toegevoegd op de plaats waar geen overlappende peptiden werden gevonden. Met een extra valine residu dat op een andere plaats werd gevonden, kwam daardoor het totale aantal aminozuren, waaruit het papaine molecule is opgebouwd, op 212. De aard van de 13 extra aminozuren werd onlangs bepaald door Lowe [80]. Het moleculgewicht van papaine, berekend uit de nu volledig bekende aminozuursamenstelling, bedraagt 23350.

De polypeptideketen is zodanig gevouwen dat het molecule in twee helften wordt verdeeld, waartussen een gleuf wordt open gelaten. In die gleuf bevindt zich het actieve centrum van het enzyme met als voornaamste bestanddelen een cysteine en een histidine residu. Deze residuen bevinden zich tegenover elkaar aan weerszijden van de gleuf, waarbij de afstand tussen hun zijketens ongeveer 4 \AA bedraagt.

In hoofdstuk I wordt een kort overzicht gegeven van de talrijke chemische onderzoeken die aan papaine zijn gedaan.

Papaine behoort tot de klasse enzymen die een vrije thiol (SH) groep nodig hebben voor hun activiteit. Het enzyme werkt als katalysator bij de hydrolytische splitsing van bepaalde peptiden, zuoramiden en esters. Papaine is weinig specifiek daar het peptidebindingen, gevormd door bijna alle aminozuren, kan hydrolyseren. Een zekere voorkeur voor bindingen van het type Lys-X of Arg-X werd tot nu toe aangenomen [12], maar kortgeleden vonden Schechter en Berger [13] een veel sterkere voorkeur voor splitsing van de X-Y binding in peptiden, waarin phenylalanine voorafgaat aan willekeurige aminozuurresiduen X en Y (-Phe-X \downarrow Y-).

In papaine, bereid volgens de methode van Kimmel en Smith [9], komt de SH-groep vrijwel niet als zodanig voor. Van een deel van de moleculen heeft de SH-groep gereageerd met cysteine, onder vorming van een gemengd disulfide [26, 62], van een ander deel is de SH-groep irreversibel gemodificeerd, waarschijnlijk door oxidatie. Het enzyme kan worden geactiveerd door omzetting van de disulfide groep en de vrije SH-groep, b. v. door een overmaat cysteine. Bovendien dient men een metaalbindend reagens, zoals EDTA, toe te voegen omdat papaine sterk wordt geremd door sporen van allerlei metalen.

In Hoofdstuk II worden de verschillende kristalmodificaties van papaine beschreven (Tabel IV). De orthorhombische vorm C, gekristalliseerd uit 70% methanol, werd gebruikt voor de structuurbepaling. Het enzyme is in dit milieu niet actief. Op verschillende manieren is echter aangetoond [51, 118, 136] dat de conformatie van papaine in het kristallisatiemedium niet significant verschilt van de natieve conformatie.

De fasehoeken van de reflecties werden bepaald met behulp van de isomorfe vervangingsmethode. Daartoe werd een serie, zware atomen bevattende, isomorfe derivaten van papaine bereid. Door het meten van de intensiteitsverschillen van corresponderende reflecties konden de fasehoeken worden berekend. Vijf derivaten, met PCMB, PCMS, PCMA, (Pap-S)-HgCl en (Pap-S)-HgCl + HgCl₂ werden gebruikt (Tabel V). Bovendien werd de anomale verstrooiing van de Hg-atomen in de derivaten met succes toegepast bij de fasebepaling.

In dit hoofdstuk wordt verder beschreven hoe de intensiteiten van de afgebogen röntgenstralen werden gemeten. Voor

de structuurbepaling met laag oplossend vermogen was dit gedaan door het opnemen van precessiefoto's. In het in dit proefschrift beschreven onderzoek werd gebruik gemaakt van een automatische diffractometer. Enkele karakteristieke eigenschappen van de door Arndt en Phillips ontworpen lineaire diffractometer [126] worden beschreven, evenals van de met drie tellers uitgeruste versie van dit instrument [127, 128].

Verder worden besproken de semi-empirische absorptiecorrecties die werden toegepast en de schaling van de intensiteiten, afkomstig van verschillende kristallen en van verschillende derivaten.

In Hoofdstuk III wordt vermeld hoe de fasehoeken werden bepaald uit de intensiteitsgegevens. De gemiddelde "figure of merit" [151] die tenslotte werd gevonden voor 9525 reflecties, was 0,82. Dit is een hoge waarde, vergeleken met andere eiwitstructuurbepalingen. De intensiteitsverschillen die optreden tussen Bijvoetparen van reflecties, als gevolg van de anomale verstrooiing door de Hg-atomen, bleken een belangrijke invloed te hebben op de berekende fasehoeken, zodat we besloten deze gegevens te benutten bij de faseberekening.

In hoofdstuk IV wordt de interpretatie van de "beste" [150] electronendichtheidsverdeling besproken. Praktisch alle zijgroepen zijn duidelijk zichtbaar als gebieden van hoge dichtheid, die zijwaarts uit de continu doorlopende dichtheid van de hoofdketen steken. Het oplossend vermogen bleek voldoende om zijgroepen met specifieke kenmerken, zoals tryptofaan, tyrosine, cysteine, isoleucine etc. te herkennen. In de hoofdketen kunnen de meeste carbonylgroepen geïdentificeerd worden doordat ze de relatief hoogste electronendichtheid bezitten. In het papaine molecuule worden vijf stukken rechtsdraaiende α -helix waargenomen, die tezamen ongeveer 60 residuen omvatten, dat is 28% van het totale aantal. Dit stemt goed overeen met het α -helix gehalte voorspeld op grond van recente ORD-metingen [201]. Ongeveer 32 residuen komen voor in een z.g. "twisted pleated sheet"-structuur (fig. 24). Slechts een tiental hiervan vormen een regelmatige antiparallelle "pleated sheet" of β -structuur [177].

De belangrijkste elementen van het actieve centrum van

papaine worden beschreven in § 4.3. Tevens wordt aandacht besteed aan de mechanismen van de katalytische werking van papaine, zoals die kortgeleden zijn voorgesteld [37, 185].

Tenslotte worden de resultaten besproken van enkele verschilfouriersyntheses van gemodificeerde papaine t.o.v. het natieve eiwit. In de eerste plaats die van het door Wolthers [189] bereide derivaat van papaine waarin de SH-groep heeft gereageerd met tosyl-lysyl-chloorketon (TLCK). In de tweede plaats werden bepaalde aminogroepen van het eiwit gemerkt door reactie met het Edman-reagens phenylisothiocyanaat en met de geïodeerde vorm van deze verbinding. De posities van de zwavel- en jodiumatomen konden worden bepaald. Het bleek dat 3 lysines en de α -NH₂ groep van het eiwit gereageerd hadden. In de derde plaats werden verschilfouriers berekend (voor de [001] en [100] projectie), van een papainederivaat waarin een aantal tyrosines waren geïodeerd met KI₃. Tenminste 4 van de 7 gevonden maxima kunnen worden gecorreleerd met tyrosineresiduen.

Deze resultaten bevestigen nog eens de juistheid van de aminozuurvolgorde, zoals die door ons uit de electronendichtheidsverdeling werd bepaald (fig. 21).