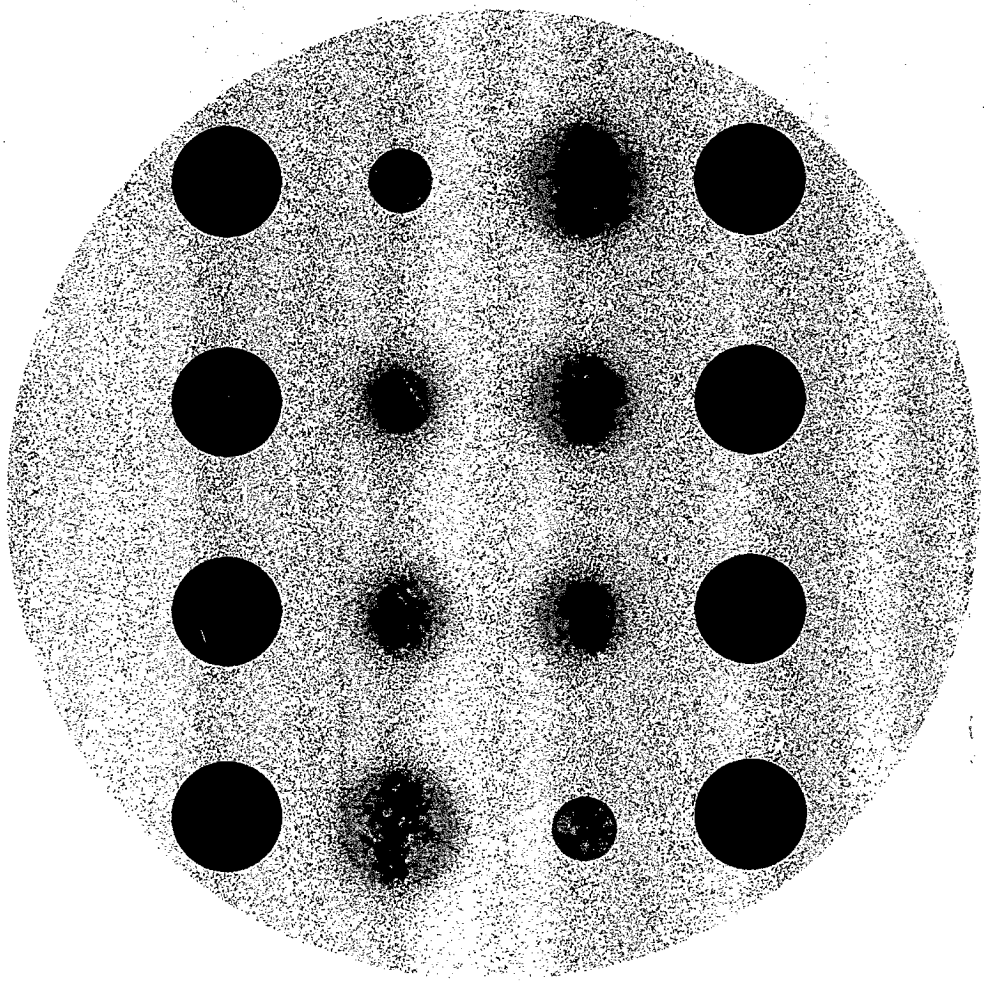
CHITIN HANDBOOK

Edited by
Riccardo A. A. Muzzarelli
and
Martin G. Peter



European
Chitin
Society

Chitin Handbook, R.A.A. Muzzarelli and M.G. Peter, eds.,
European Chitin Society. 1997. ISBN 88-86889-01-1

3D-Structure and enzymatic mechanism of chitinase A from the chitinolytic soil bacterium *Serratia marcescens*

Constantinos E. Vorgias[a] and Kyriacos Petratos[b]

[a]Athens University, Biology Department, Div. of Biochemistry and Molecular Biology,
Biochemistry Laboratory, Panepistimiopolis-Kouponia, 15701 Athens, Greece
Fax: +30-1-7257572, email: cvorgias@biology.db.uoa.gr
[b]IMBB-FORTH, P.O. Box 1527, 71110 Heraklion, Crete, Greece
Fax:+30-81-394406, email: petratos@imbb.forth.gr

X-ray diffraction (or X-ray crystallography) is one of the most powerful methods for 3D structure analysis of molecules with molecular masses up to several 100 kDa (Blundell and Johnsson, 1976; Drenth, 1994). First, the symmetry of the crystal lattice (space group) and the parameters of the unit cell (dimensions of a, b, c, and values of the angles α, β, γ) are determined. A few "still" X-ray photographs are taken of a crystal at different fixed positions. Then, some small angle (0.5°) photographs are taken with the crystal rotating slowly around an axis perpendicular to the direction of the X-rays. During characterization, the diffraction limits (e.g. 1.5 Å) and quality of the lattice (single or twin and mosaicity) are determined. The crystal content (as percent by volume) of mother liquor is estimated [usually 40 - 70 % (v/v)] also at this stage.

If the crystal is single (one lattice throughout the volume of the crystal) and has a reasonable mosaicity (ca. 0.3°), i.e. limited disorder of the "ideal" blocks forming the entire crystal, the intensities of the several unique beams diffracted from the crystal can be measured. For a typical protein crystal the intensities to be measured are usually of the order of 10^3 - 10^5. These independent (unique) data are computer processed to yield structure factor amplitudes (F's).

During an X-ray diffraction experiment, no phases of the diffracted beams can be measured. In order to calculate the values of the phase (a) for each reflection separately, isomorphous crystals containing selected heavy atom compounds are produced and their diffraction patterns are compared with that of the native crystal. This method is known as multiple isomorphous replacement (MIR). In case we know the structure of a homologous (similar) molecule, we

employ the method of molecular replacement (MR) in order to obtain a starting set of phase angles. Another method, multiple wavelength anomalous diffraction (MAD), is used increasingly at synchrotron sites. In this method, advantage is taken of the anomalous scattering components of certain atoms which are either contained in the native structure or are inserted (e.g. Se) by genetic techniques (growing auxotrophic strains in Se-Met containing medium). Finally, the direct methods of phase calculation can apply only for relatively small structures consisting of up to about 100 non-hydrogen atoms. Nevertheless, recent developments may lead to a breakthrough in phase determination for macromolecules.

When the phases are known, we calculate an electron density map and display it on the monitor of a computer graphics system. The interpretation of the electron density map consists of constructing a molecular model. By using higher resolution data, we can refine the structure, i.e. calculate better and more accurate phases for the Bragg reflections which in turn provide more accurate coordinates for the atoms of the structure.

For molecules with masses up to ca. 25 kDa, complementary structural information can be provided by various NMR techniques.

Materials and facilities for crystallization and crystal mounting

A comprehensive handbook which describes the crystallization techniques as well as the entire methodoly was published (Ducruix and Giegé, 1992). Briefly, the following material and facilities are required to carry out crystallization experiments:

We start crystallization attempts with a homogeneous sample of the molecule under examination (e.g. enzyme). This must be done in a crystallization room under strict temperature, humidity and vibration control! For the vapour diffusion hanging or sitting drop technique, various trays (plates) are used. Highly pure buffer solutions and precipitating agents are required. The crystal growth is monitored with a stereo microscope (zooming option is advisable). A thin long needle and thin-walled capillaries of varying diameter (e.g. 0.5 - 1.5 mm, depending on the crystal size) are necessary to handle the crystal to be mounted. Syringes and teflon tubing are used to remove liquid which surrounds the crystals. Excess mother liquor is removed by means of filter paper. The ends of the capillary carrying the crystal are sealed with low melting wax in order to prevent the crystal from drying and disintegrating.

Facilities for X-ray data collection

The home laboratory is equipped with a medium resolution (ca. 2.8 Å) X-ray instrument (rotating copper anode generator) for data collection from native

protein crystals and heavy atom derivatives. The generator should be equipped with a pair of focusing mirrors to monochromatize and improve the intensity of the primary X-ray beam impinging on the crystal. A synchrotron radiation source for data collection at high resolution (< 2.0 Å) is accessible at various sites.

Recording of the diffracted X-rays from the crystal is achieved by means of an efficient area detector (e.g. image plate scanner) which should provide good spatial resolution characteristics (ca. 1 mm) and large dynamic range (10^5). Suitable computer hard- and software is required for control of the detector, processing the raw data, calculating the electron density maps, fitting a model in the map and refinement of the resulting structure.

Measurements of protein crystals

Since the Bragg reflections are the only data in a crystallographic analysis of a structure, the diffraction data collection both for the native and the isomorphous heavy atom derivative crystals must be done with great care.

Nearly all of the reflections within the selected resolution range should be measured, i.e. their number should be the same as that expected from the unit cell dimensions (completion 100 %). Most of the diffraction "spots" should have intensity $I > 3\ \sigma_m$, where σ_m is the standard deviation of the intensity. The values of the overall reliability factors (R_{sym}, R_{merge}) should be as small as possible (usually 3 - 4 %). These factors provide an estimation of the random and systematic errors of the measurements of the symmetry equivalent reflections or multiples recorded on different images. The values of the corresponding factors (R_{sym}, R_{merge}) for the outer resolution shell could be ca. 15 %.

Data processing and scaling

Various computer programs have been developed over the last decades for processing the raw diffraction data (raw I's) in order to obtain structure factor amplitudes (F's). Some of them are MOSFLM (Leslie *et al.*, 1986), XDS (Kabsch, 1993), and DENZO (Otwinowski, 1991). Background scattering from air and surrounding liquid, radiation damage, absorption effects, radiation polarization, geometry of the data acquisition technique, crystal volume, primary beam intensity e.t.c. are taken into consideration in order to "correct" the measured intensity.

A common situation in the crystallography of macromolecules is the availability of more than one set of data, e.g. from measurements of different crystals (native or derivatives). In this case, data processing should allow to adjust to the same relative scale of intensity. This procedure is called "scaling". The structure factor amplitudes (F's) are the square root of the "corrected" scaled intensities of diffraction.

Model building (interpretation of the electron density map)

We can claim that the structure is "solved" (determined) when the structure factor amplitudes (F's) are derived from the measured intensities and their corresponding phases (a) are calculated with a certain degree of accuracy with one of the methods mentioned above (MIR, MR or MAD). There are methods to evaluate the accuracy of the calculated phases of the diffracted X-rays, a description of which would be beyond the scope of this book.

The actual result of the crystallographic analysis is the electron density map calculated by Fourier transformation of the structure factors. In order to obtain the desired model, this computer displayed electron density map must be interpreted. This is done in a dark room. The time required for the map interpretation is inversely proportional to the quality of the measured data and the resolution of the native data (for an "average" protein usually about 2.5 Å).

Structure refinement

The term structure refinement refers to the process of obtaining atomic co-ordinates more accurately. The aim is to minimize the differences between the observed (measured) amplitudes and those calculated from the model structure factor. In order to carry out a proper refinement we need to have an overdetermined problem. This means that the structure factor amplitudes required to position all the atoms in the structure must have been measured with a multiplicity of 10. One atom has four parameters, i.e. three Cartesian co-ordinates and an isotropic temperature factor B which is explained in the following. Therefore, we need about 40 observable unique reflections per (non-hydrogen) atom in the assymetric unit.

Indeed, an even "better" data set is collect after the first model has been obtained from interpretation of an electron density map at medium resolution (2.5 Å). This requires an intense monochromatic X-ray source of very low beam divergence, as provided by a bright synchrotron beam line. The other important factor which determines the resolution limits of the diffraction is the crystal itself. Static and dynamic disorder of the crystalline lattice are manifested together in the so-called temperature factor (B) or the mean atomic displacement factor (u). The higher the temperature factor of an atom (i), the higher the mobility of the atom around its averaged "central" position (x_i, y_i, z_i). If the diffraction data extend beyond the 3 Å range, we could in principle put them on an absolute scale, i.e. make the scattering factors of the atoms directly proportional to the diffracting electrons of the unit cell. This allows to estimate the overall temperature factor of the data, which is required to estimate the limit of diffraction of the crystal.

A typical protein crystal usually diffracts weakly at about 1.5 Å spacing. This results in less than the required structure factor amplitudes for a proper refinement (ca. 5 x atomic parameters) Nevertheless, there are enough data to carry out a structure refinement because proteins consist of peptide units whose structures are known at very high resolution. Actally, "dictionaries" for bond lengths and angles which can provide indirectly the "missing" data during a typical "restrained" least-squares refinement are available (Hendrickson, 1985). The "restrained" refinement has a convergence radius of about 1 Å, i.e. it cannot on the average shift any atom more than 1 Å away. That is why we need several sessions of manual rebuildings in front of a graphics station. Methods of molecular dynamics (MD) (Brunger *et al.*, 1987) are employed for the refinement of macromolecular structures. In particular, MD is useful when the original model has extensive mistakes, i.e. the positions of certain atoms are 6 Å away from their true values. Since this method has a high convergence radius in the order of several Å, the number of manual rebuiling sessions in the beginning of the refinement can be reduced. We usually start the refinement with MD methods and finish it with the restraint least-squares method.

Crystallization of chitinase A

Recombinant chitinase A (Vorgias, *et al.*, 1993; see also Vorgias, this volume) is crystallized using the vapour diffusion method with hanging drops (Vorgias *et al.*, 1992). The hanging drop has a final volume of 10 µl and contains 5 µl protein solution (ca. 10 mg/ml) and 5 µl reservoir buffer (18 - 22 % PEG4000, 0.15 - 0.20 M ammonium sulfate, and 0.5 - 1 % 2-propanol in 100 mM acetate buffer, pH 5.2). Crystallization is carried out at 17°C and the first crystals could appear after few days or, sometimes, after several months. For the structure determination, "freshly" formed crystals are used.

The structure of chitinase A

The X-ray structure of chitinase A from the chitinolytic bacterium *Serratia marcescens* has been solved by multiple isomorphous replacement (MIR) and refined at 2.3 Å resolution, resulting in a crystallographic R-factor of 16.2 % (Perrakis *et al.*, 1994). The structure of chitinase A consists of three domains (Fig. 1).

The N-terminal domain: The amino-terminal domain (aminoacid residues 24 - 137) comprises only β-strands. This fold is similar to the immunoglobulin-like β sandwich fold and it resembles the fibronectin III (FnIII) module domain. It was also found in some other chitinases. The N-terminal domain makes a

number of hydrogen bonds with the rest of the structure, in particular to the $(\alpha\beta)_8$-barrel.

Figure 1. Ribbon diagram illustrating the structure of chitinase A, its three domains, and the groove of the active site where the sugar ring is bound.

The hinge region: The hinge region (residues 138 - 158) connects the N-terminal domain with the rest of the structure via a 21 amino acid long chain. This region shows unusual characteristics and contains 10 charged residues. It is relatively mobile and probably allows the N-terminal domain to have various relative positions in solution.

The $(\alpha\beta)_8$ barrel: The eight stranded $(\alpha\beta)_8$ barrel is the catalytic domain of the enzyme (residues 159 - 442 and 517 - 563). It has a number of irregularities compared with a typical $(\alpha\beta)_8$ barrel fold, such as that of triose phosphate isomerase.

The third domain has an $\alpha+\beta$ fold and is formed by an insertion in the barrel motif (residues 443 - 516). It is comprised of 5 β-strands of which one is interrupted forming an all anti-parallel b-sheet. An α-helix protects the first hydrophobic surface of the sheet. The rest of the surface is protected by a 3_{10} helix together with some coil structure.

Proposed reaction mechanism

The structures of several family 18 glycosyl hydrolases have been reported, among them hevamine, a plant defense protein with combined chitinase and lysozyme activity (Terwisscha van Scheltinga *et al.*, 1994; see also Terwisscha

van Scheltinga and Dijkstra, this volume), chitinase A from *Serratia marcescens* (Perrakis *et al.*, 1994), and endo-β-*N*-acetylglucosaminidases F1 (van Roey *et al.*, 1994), and H (Rao *et al.* 1995). A common fold for all family 18 enzymes is an $(\alpha\beta)_8$ barrel (Davies and Henrissat, 1995). Substrate binding takes place at a long groove which is located twoards the C-terminal end of the $(\alpha\beta)_8$ barrel aminoacid sequence.

Enzymatic hydrolysis of the glycosidic bond takes place via general acid-base catalysis which is initiated by protonation of the glycosidic oxygen by an acidic glutamic acid residue (Sinnot, 1990; Davies and Henrissat, 1995). The positive charge developing at C-1 is a mesomeric carboxonium ion. In the case of lysozyme, it is thought that the carboxonium ion is stabilized by a nucleophilic aspartate. Attack of a water molecule from the equatorial side results in overall retention of anomeric configuration by double displacement. On the other hand, direct additon of water from the axial side (single displacement) leads to inversion of configuration. The distance between the glutamic acid and the aspartate residues is indicative for the mode and stereochemistry of the reaction: 4.8 - 5.3 Å is typical for retention and 9 - 9.6 Å for inversion. In that case, a water molecule is positioned between the anomeric carbon and the aspartate.

Biochemical experiments have clearly shown that hevamine (Terwisscha van Scheltinga *et al.*, 1995) and several bacterial chitinases (Davies and Henrissat, 1995; Armand *et al.*, 1994) are retaining enzymes. From structure comparisons and mutagenesis experiments, it is clear that the proton donor is a glutamate residue (127 in hevamine, 315 in chitinase A and 132 in endo-β-*N*-acetylglucosaminidases F1 and H). Its acidity is enhanced by a strong hydrogen bond of an aspartate in position (n-2). In all three structures no negatively charged residue can be unambiguously assigned to stabilize a carboxonium ion intermediate. A mechanism involving stabilization by interaction with the O-7 carbonyl oxygen of the *N*-acetyl group from the axial (α) side, i.e. substrate assisted catalysis, is also strongly supported by crystallographic data of complexes of the chitinase inhibitor allosamidin with hevamine (Terwisscha van Scheltinga *et al.*, 1995) and with chitinase A (unpublished observations). Finally, in chitobiase, the *N*-acetyl group seems to be deformed and is in close proximity to the C_1 atom (Tews *et al.*, 1996a, 1996b).

This mechanism is in contrast to many other retaining glycosyl hydrolases which stabilize an intermediary carboxonium intermeditae by a nucleophilic aspartate residue.

Mode of interaction of chitinases with substrates and inhibitors

The structures of hevamine with $(NAG)_3$ and allosamidin (Terwisscha van Scheltinga *et al.* 1995), and of chitinase A with $(NAG)_2$ and allosamidin (unpublished results) reveal numerous differences as well as striking similarities in the binding of the sugars to these proteins. These enzymes, together with endo-β-*N*-acetyl-glucosaminidases F1 and H, share a common evolutionary origin. They possess a common fold but their specificities were adapted to different substrates, i.e. lysozyme activity of hevamine in contrast to chitinase A which is a more efficient enzyme for the cleavage of chitin. Several subsites have been defined for the binding of the substrate in the groove around the catalytic site.

The (+2) subsite in chitinase A is mainly defined by the stacking of the aromatic residue Phe396 against the hydrophobic face of the sugar. Two direct hydrogen bonds are present (i) between the carboxylate oxygen of Asp391 and N_2 of the *N*-acetyl group of the sugar, and (ii) between the Nζ atom of Lys320 and O-8 of the sugar. The Phe396 ring is absent in hevamine and no sugar is bound in this subsite. This explains the apparent higher affinity of chitinase A for chitobiose in this subsite. The (+2) site is absent in endo-β-*N*-acetyl-glucosaminidase F1, reflecting the difference in substrate specificity, since in that position a linkage with an Asn residue is expected.

The (+1) subsite in chitinase A is defined by the aromatic ring of Trp275 which is stacked against the hydrophobic face of the sugar. No direct hydrogen bonds exist and the affinity is modulated via three water molecules that mediate hydrogen bonds betwen the O-3 and O-6 atoms of the sugar and the protein. The Trp275 ring of chitinase A is not present in either of the other enzymes. This supports the assumed higher affinity of this site in chitinase A as compared with hevamine. Upon cleavage of peptigoglycans by hevamine, a *N*-acetylmuramate (NAM) residue is expected to occupy that site. There are no obvious structural reasons why a NAM residue could not be accomodated at that position in chitinase A. The fact that this ring can form more favourable hydrogen bonds in endo-β-*N*-acetyl-glucosaminidase F1 reflects the need for the tight binding of this sugar, since this is the only interaction for the docking of the 'product' part of the substrate.

The distance between the O-4 atom of the sugar and the carboxyl oxygen of the catalytic residue Glu315 is 3.2 Å in chitinase A. This distance is longer than expected for nucleophilic attack. Upon binding of the rest of the substrate at the (-) sites, a slight movement of the sugars bound in the (+) sites is to be expected without influencing the hydrophobic packing. It can lead to the formation of more direct hydrogen bonds to this sugar ring.

Subsite (-1) contains an aromatic ring of Trp539 which is stacked against the hydrophobic face of the allosamizoline moiety of allosamidin as revealed by the structures of inhibitor complexes of chitinase A and hevamine. A conserved Trp residue is present also in lysozyme and in endo-β-*N*-acetyl-glucosaminidase F1.

Subsite (-2) forms hydrogen bonds between the carboxyl oxygens of Glu540 and Glu473 and O-7 and N-2 atoms of the *N*-acetyl group of the second sugar, respectively. The O-6 atom of the sugar is linked via a hydrogen bond with the main chain nitrogen of Thr276.

The (-3) subsite is defined in chitinase A by Trp167, which packs almost perfectly against the sugar ring, as compared with hevamine where that Trp is absent. The two direct hydrogen bonds in hevamine are disrupted and presumably substituted by the hydrophobic interaction. This is a very interesting subsite since in chitinase A it should only bind NAG, in hevamine either NAG or NAM and in endo-β-*N*-acetyl-glucosaminidase F1 a mannose.

The (-4) subsite of chitinase A binds the sugar in a different manner than in hevamine. On moving further from the active site it might be expected that the mode of binding is less well conserved and the sugar ring could adopt a different position. This site is completely lost in endo-β-*N*-acetyl-glucosaminidase F1, reflecting the difference in substrate specificity. Of interest is the 'exchange' of the two direct hydrogen bonds in hevamine for the hydrophobic interaction with Trp in chitinase A.

From our modelling studies on chitinase A with allosamidin, it is possible that there is one more binding site (-5) which does not exist in hevamine or endo-β-*N*-acetyl-glucosaminidases F1 or H.

References

Armand S, Tomita H, Heyraud A, Gey G, Watanabe T and Henrissat B (1994). FEBS Lett. 343: 177-180.

Blundell TL and Johnson LN (1976). *Protein Crystallography*, Academic Press, London.

Brunger AT, Kuriyan J and Karplus M (1987). Science 235: 458-460.

Davies G and Henrissat B (1995). Structure 3: 853-859.

Drenth J (1994). *Principles of Protein X-ray Crystallography*, Springer-Verlag. New York.

Ducruix A and Giegé R (1992). *Crystallization of Nucleic Acids and Proteins, a Practical Approach*, IRL Press, Oxford.

Hendrickson WA (1985). Meth. Enzymol. 115: 252-270.

Kabsch W (1993). J. Appl. Cryst. 26: 795-800.

Leslie AGW, Brick P and Wonacott AJ (1986). CCP4 News 18: 33-39.
Otwinowski Z (1991). DENZO. A Film Processing Program for Macromolecular Crystallography. Yale University, New Haven..
Perrakis A, Tews I, Dauter Z, Oppenheim A, Chet I, Wilson KS and Vorgias CE (1994). Structure 2: 1169-1180.
Perrakis A, Ouzounis C, Wilson KS and Vorgias CE (1996). Advan. Chitin Science 1: 34-41.
Rao V, Guan C and Roey PV (1995). Structure 3: 449-457.
Sinnot ML (1990). Chem. Rev. 90: 1171-1202.
Terwisscha van Scheltinga AC, Armand S, Kalk KH, Akira I, Henrissat B and Dijkstra BW (1995). Biochemistry 34: 15619-15623.
Terwisscha van Scheltinga AC, Kalk KH, Beintema JJ and Dijkstra BW (1994). Structure 2: 1181-1189.
Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS and Vorgias CE (1996a). Nature Struct. Biol. 3: 638-647
Tews I, Wilson KS and Vorgias CE (1996b). Advan. Chitin Science 1: 26-33.
van Roey PV, Rao V, Plummer TH and Tarentino A (1994). Biochemistry 33: 13989-13996.
Vorgias CE, Kingswell AJ, Dauter Z and Oppenheim BA (1992). J. Mol. Biol. 226: 897-898.
Vorgias CE, Tews I, Perrakis A, Wilson KS and Oppenheim BA (1993). In: Muzzarelli RAA (ed.). *Chitin Enzymology*, Atec, Grottamare, pp. 417-422.