

# The $\alpha$ -amylase from the yellow meal worm: complete primary structure, crystallization and preliminary X-ray analysis

Stefan Strobl<sup>a</sup>, Franz-Xaver Gomis-Rüth<sup>1,b</sup>, Klaus Maskos<sup>a,b</sup>, Gerhard Frank<sup>a</sup>,  
Robert Huber<sup>b</sup>, Rudi Glockshuber<sup>a,\*</sup>

<sup>a</sup>Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg, CH-8093 Zürich, Switzerland

<sup>b</sup>Max-Planck-Institut für Biochemie, D-85152 Planegg-Martinsried, Germany

Received 10 April 1997

**Abstract** The  $\alpha$ -amylase from *Tenebrio molitor* larvae (TMA) was purified from a crude larval extract. After removal of the N-terminal pyroglutamate residue and identification of the following 17 residues by Edman sequencing, the cDNA of mature TMA was cloned from larval mRNA. The encoded enzyme consists of 471 amino acid residues and has 57–79% sequence identity to other insect  $\alpha$ -amylases and also shows high homology to the mammalian enzymes. TMA was crystallized in form of well-ordered orthorhombic crystals of space group  $P2_12_12_1$  diffracting beyond 1.6 Å resolution with unit cell dimensions of  $a = 51.24$  Å,  $b = 93.46$  Å,  $c = 96.95$  Å. TMA may serve as model system for the future analysis of interactions between insect  $\alpha$ -amylase and proteinaceous plant inhibitors on the molecular level.

© 1997 Federation of European Biochemical Societies.

**Key words:** *Tenebrio molitor*;  $\alpha$ -Amylase; Protein sequence; Crystallization

## 1. Introduction

$\alpha$ -Amylases ( $\alpha$ -1,4-glucan-4-glucanohydrolases, EC 3.2.1.1) constitute a family of enzymes that catalyze the hydrolysis of  $\alpha$ -D-(1,4)-glucan linkages in starch and related compounds. Therefore, they play a central role in carbohydrate metabolism. Living on a polysaccharide-rich diet, many organisms depend on the effectiveness of their amylases for survival. This is particularly true for insects like the meal beetle *Tenebrio molitor*, a cosmopolitan pest of grain products. Its larva contains a single  $\alpha$ -amylase (TMA) that is an acidic protein with a pH optimum for the cleavage of starch of 5.8 [1]. The enzyme is accordingly well adapted to its physiological environment in the larval midgut, where a slightly acidic pH is prevalent [2]. The catalytic properties of affinity-purified TMA with the substrate starch have been reported before [1]. In addition, it was shown that TMA is inhibited by proteinaceous plant inhibitors [3–9]. Here we describe the purification of the enzyme and the complete amino acid sequence obtained from protein sequencing and the cDNA sequence. Furthermore, we report the crystallization of TMA and discuss the differences in primary structure between mammalian and insect  $\alpha$ -amylases.

## 2. Materials and methods

### 2.1. Protein purification

Yellow mealworms (larvae of *Tenebrio molitor*) were bought at a local pet shop. Larvae were allowed to pupate and imagoes were used to verify the species. TMA was obtained from a crude extract of 250 g larvae by a 5-step purification method. Yellow mealworms were homogenized with a blender in 750 ml of standard buffer (20 mM acetic acid/NaOH (pH 5.4), 1 mM CaCl<sub>2</sub>). The homogenate was centrifuged (23 000  $\times g$ , 30 min, 4°C) and ammonium sulfate was added to the supernatant to 90% saturation. The precipitated protein was collected by centrifugation (48 000  $\times g$ , 30 min, 4°C), suspended in 100 ml of standard buffer and extensively dialyzed against the same buffer. After centrifugation (48 000  $\times g$ , 30 min, 4°C), the supernatant was applied to a DE-52 anion exchange column (45 ml, Whatman) equilibrated with standard buffer. A linear NaCl gradient (0–500 mM) was used to elute the proteins. Fractions with  $\alpha$ -amylase activity were pooled, dialyzed against standard buffer containing 1 M ammonium sulfate, and applied on a Phenyl Sepharose HP column (45 ml, Pharmacia). Proteins were eluted with a linear ammonium sulfate gradient (1.0–0 M). Fractions with  $\alpha$ -amylase activity were combined, dialyzed against standard buffer and applied to a Resource Q anion exchange column (6 ml, Pharmacia). Proteins were eluted with a linear NaCl gradient (0–500 mM). Fractions with  $\alpha$ -amylase activity were pooled and concentrated to a volume of 2 ml. The sample was applied to a Superdex 200 HighLoad 26/60 gel filtration column (Pharmacia) and eluted with standard buffer. The fractions containing pure TMA could be stored at 4°C without any detectable degradation for at least 9 months. Typically, 25 mg homogeneous TMA were obtained by this procedure.

### 2.2. Determination of the N-terminal sequence of TMA

Pyroglutamate aminopeptidase (Boehringer Mannheim) treatment made the blocked N-terminus of TMA accessible for Edman degradation. Approximately 80  $\mu$ g TMA were denatured by heating for 30 min at 80°C in 11  $\mu$ l of standard buffer containing 7.4 M guanidinium chloride and 90 mM dithiothreitol (DTT). After addition of 29  $\mu$ l of distilled water, 40  $\mu$ l of 200 mM sodium phosphate (pH 8.0), 20 mM EDTA, 10 mM DTT, 10% (w/v) glycerol, and 1.25  $\mu$ g pyroglutamate aminopeptidase, the sample was incubated at room temperature for 5 h, again heated to 80°C for 30 min and vacuum-dried. After addition of distilled water to the previous volume and 1.25  $\mu$ g pyroglutamate aminopeptidase, the reaction was performed as described above. Overall, the reaction was repeated six times after heat denaturation and drying.

One-fifth of the sample was subjected to SDS-polyacrylamide gel electrophoresis [10] and blotted onto a polyvinylidene difluoride membrane [11]. The band corresponding to TMA was N-terminally sequenced on a Modular Sequencer (Dr. Ing. H. Knauer GmbH, Berlin) [12] which had been modified to allow isocratic identification of the phenylthiohydantoin amino acids [13].

### 2.3. cDNA synthesis, PCR amplification and DNA sequencing

Total RNA from a last instar larva of *T. molitor* was isolated with the 'Oligotex Direct mRNA kit' (Qiagen). First-strand synthesis was carried out with the 'cDNA cycle kit' (Invitrogen). Both kits were used according to the manufacturer's instructions.

The following deoxyoligonucleotide primers were used for the amplification of TMA cDNA fragments by the polymerase chain reac-

\*Corresponding author. Fax: (41) 1-633-1036.  
E-mail: RUDI@MOL.BIOL.ETHZ.CH

<sup>1</sup>Present address: Centre d'Investigacions i Desenvolupament C.S.I.C., Jordi Girona, 18–26, E-08034 Barcelona, Spain.



```

1  CAGAAGGACCGAATTTTTCGAAGTGGTAGAAATAGCATCGTGCACTTGTTCGAATGGAAATGGAATGACATC 72
1  Q K D A N F A S G R N S I V H L F E W K W N D I 24

73  GCCGACGAATGCCGAGAGATTCTTCGACCCCAAGGATTTCGGAGGAGTTCAGATCTCTCCACCTAACGAGTAC 144
25  A D E C E R F L Q P Q G F G G V Q I S P P N E Y 48

145  TTGGTGGCGGATGGCAGACCCCTGGTGGGAACGGTACCAACCCGTGAGCTACATCATCAACACCAAGGTCTGGA 216
49  L V A D G R P W W E R Y Q P V S Y I I N T R S G 72

217  GACGAATCGGCCTTCACTGACATGACCAGACGCTGCAACGATGCTGGTGTTCGTATTTATGTGGATGCTGTG 288
73  D E S A F T D M T R R C N D A G V R I Y V D A V 96

289  ATCAACCACATGACTGGAATGAACGGGGTCGGTACCTCTGGAAGCTCAGCTGATCACGACGGCATGAATTAT 360
97  I N H M T G M N G V G T S G S S A D H D G M N Y 120

361  CCAGCTGTACCGTATGGTTCGGGAGATTTCACAGCCCTTGTGAAGTCAACAACCTACCAAGACGCTGACAAC 432
121  P A V P Y G S G D F H S P C E V N N Y Q D A D N 144

433  GTGAGGAACCTCGAACTTGTAGGTCTTCGAGATTGGAATCAGGGGTCAGATTATGTGAGGGGCGTGCTCATC 504
145  V R N C E L V G L R D L N Q G S D Y V R G V L I 168

505  GACTACATGAACCATATGATCGATTTCGGGGTGGCTGGATTGAGTGGATGCCGCCAACACATGTCCGCT 576
169  D Y M N H M I D L G V A G F R V D A A K H M S P 192

577  GGAGATCTGAGTGTGATCTTCTCCGGCTTGAAAAATTTGAACACCGATTACGGCTTCGCAGACGGCGCTAGA 648
193  G D L S V I F S G L K N L N T D Y G F A D G A R 216

649  CCCCTCATCTACCAAGAAGTTATAGATCTGGGTGGTGGGCTATCAGCAAGAACGAGTACACAGGCTTTGGT 720
217  P F I Y Q E V I D L G G E A I S K N E Y T G F G 240

721  TCGTCTTGGAAATTCAGTTCGGAGTCAGTCTAGGCAACGCCCTTCAGGGTGGAAACCAAGTTGAAGAAATTTG 792
241  C V L E F Q F G V S L G N A F Q G G N Q L K N L 264

793  GCGAATCGGGTCCAGAATGGGGTCTACTCGAAGGCCTAGACGCTGTGTGTTCGTCGACAATCACGACAAT 864
265  A N W G P E W G L L E G L D A V V F V D N H D N 288

865  CAACGTACCGCGGGGAGTCAAATTTTGACGTACAAGAACCCCAAGCCGTACAAAATGGCGATCGCTTTCATG 936
289  Q R T G G S Q I L T Y K N P K P Y K M A I A F M 312

937  TTGGCCCATCCTTATGGGCACCACAAGGATCATGTCCAGTTTGTACTTCACCGACAACGATCAAGGACCTCCT 1008
313  L A H P Y G T T R I M S S F D F T D N D Q G P P 336

1009  CAAGATGGCAGCGGCAACTTGATTTCCTCGGAATCAATGACGACAACACCTGTAGCAATGGATACGCTCGC 1080
337  Q D G S G N L I S P G I N D D N T C S N G Y V C 360

1081  GAGCACCGTTGGAGGCGAGTTTACCGGAATGGTGGGATTTCAGAAATGCGGTTCGAAGGACACAAGTAGAGAAT 1152
361  E H R W R Q V Y G M V G F R N A V E G T Q V E N 384

1153  TGGTGTTCGAATGATGACAACCAGATCGCCTTCAGTCGAGGAAGTCAAGGATTGTAGCGTTTACCAACGGT 1224
385  W S N D D N Q I A F S R G S Q G F V A F T N C 408

1225  GGAGACTTGAACCAAAACCTCAACACTGGACTTCCTGCTGGTACTTATTCGACGTTATCTCCGGAGAGTTG 1296
409  G D L N Q N L N T G L P A G T Y C D V I S G E L 432

1297  TCCGGTGGGTCTTGCACCGGCAATCTGTAAACAGTTGGAGATAACGGATCTGCTGATATTCTTTGGGAAGT 1368
433  S G G S C T G K S V T V G D N G S A D I S L G S 456

1369  GCCGAAGATGATGGAGTCTAGCTATCCATGTTAACGCAAAATGTAAATAATGTAATAAGACGATTCGAA 1440
457  A E D D G V L A I H V N A K L * 471

1441  CAGT - poly A tail 1444

```

Fig. 2. TMA cDNA sequence and deduced protein sequence of mature TMA. The first 20 nucleotides coding for residues 1–7 correspond to the synthetic deoxynucleotide primer which was used for amplification of the gene by PCR and which had been deduced from N-terminal Edman sequencing.

lases from *Tribolium castaneum* and *Drosophila pseudoobscura*.

Mature TMA consists of 471 amino acids, has a molecular mass of 51.3 kDa and a calculated pI of 4.3. The primary structure of TMA is more than 57% identical to the known  $\alpha$ -amylase sequences from insects and exhibits the same characteristic differences from the mammalian enzymes as the other insect  $\alpha$ -amylases (Fig. 3).

### 3.3. Determination of catalytic parameters of TMA

Catalytic parameters of TMA were determined with the disaccharide substrate *p*-nitrophenyl- $\alpha$ -D-maltoside (NPM) at 25°C and pH 5.8, which is the pH optimum for cleavage of starch by TMA [1]. NPM is only a poor  $\alpha$ -amylase substrate,

but allows on-line detection and quantification of substrate hydrolysis [21]. The  $K_M$  value for the cleavage of NPM was found to be  $43 \pm 7$  mM at 25°C and pH 5.8 and  $k_{cat}$  was determined to be  $0.83 \text{ s}^{-1}$ .

### 3.4. Crystallization of TMA and preliminary X-ray analysis

TMA crystals appeared after a few days and grew to a maximum size of  $1.5 \times 0.4 \times 0.4$  mm within 2 weeks (Fig. 4). They belong to the orthorhombic space group  $P2_12_12_1$  with cell constants  $a = 51.24 \text{ \AA}$ ,  $b = 93.46 \text{ \AA}$ ,  $c = 96.95 \text{ \AA}$ , and have one molecule per asymmetric unit. The tight packing of the molecules within the crystal is reflected by a Matthews-parameter of  $2.24 \text{ \AA}^3/\text{Da}$  (corresponding to 45% solvent content), explaining the high diffraction power up to more than  $1.6 \text{ \AA}$

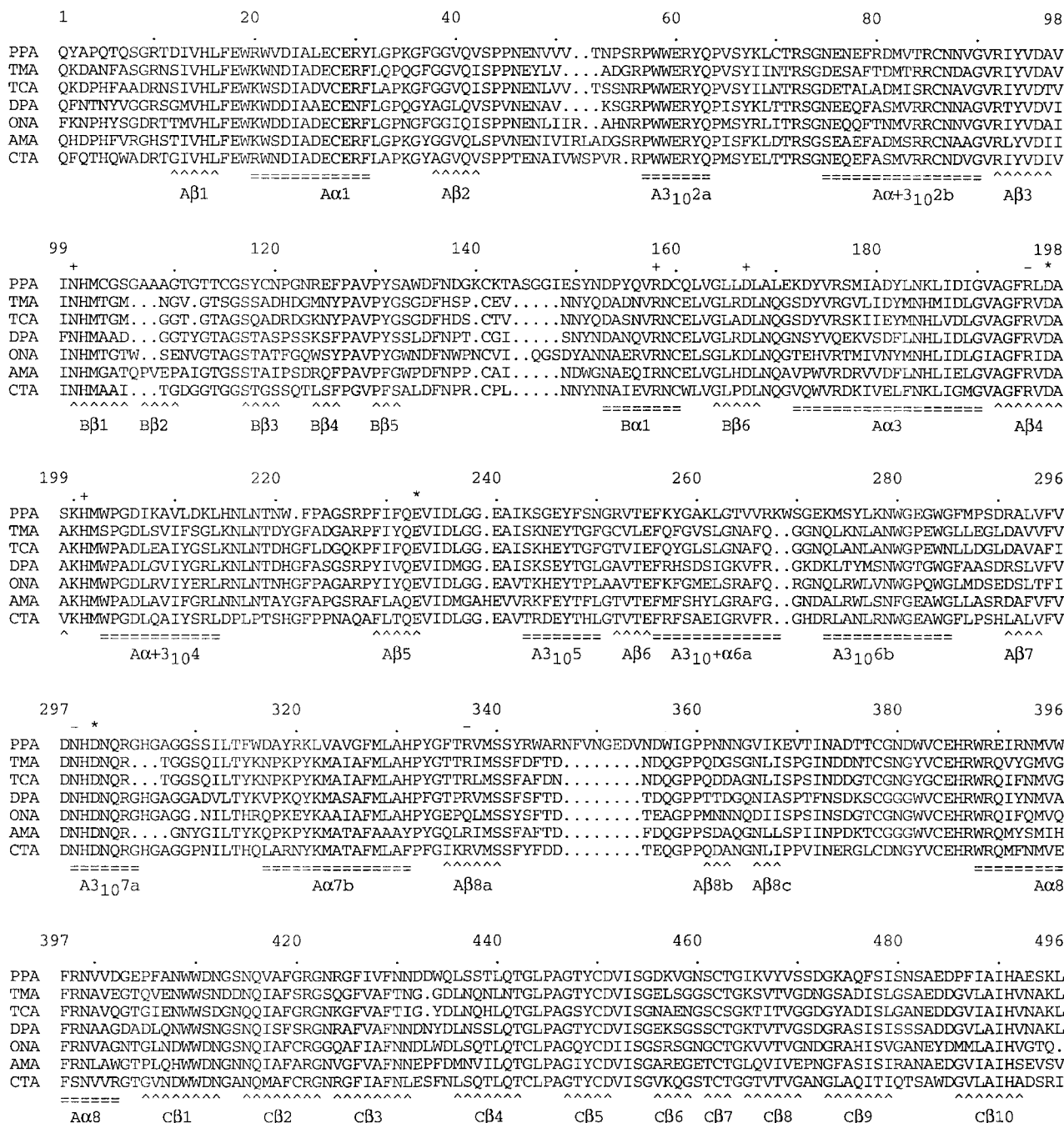


Fig. 3. Alignment of TMA's amino acid sequence with the α-amylase sequences from pig pancreas and from five insect species (amino acid numbering according to PPA). Amino acids involved in catalysis are depicted by asterisks, and those co-ordinating Ca<sup>2+</sup> and Cl<sup>-</sup> by (+) and (-), respectively. Secondary structure elements of PPA are depicted by (=) for helices and by (-) for β-strands [23]. The accession codes for the α-amylase sequences and their identities/similarities with TMA are as follows: PPA-pig pancreatic α-amylase (P1r1:alpgp) [28], 53.8%/71.5%; TMA-*Tenebrio molitor* α-amylase (PIR2: S75702); TCA-*Tribolium castaneum* α-amylase (em-in:u04271\_B) [29], 78.6%/87.5%; DPA-*Drosophila pseudoobscura* α-amylase (em-in:u20335) [30], 66.3%/77.7%; ONA-*Ostrinia nubilalis* α-amylase (em-in:u04223) [31], 61.6%/78.5%; AMA-*Anopheles merus* α-amylase (patchX:u01210) [32], 60.0%/75.5%; CTA-*Culex tarsalis* α-amylase (patchX:u01211) [33], 56.7%/72.8%.

resolution with conventional CuK<sub>α</sub> radiation. A total of 244244 measurements was collected and processed to 1.64 Å resolution which were merged into 58219 independent reflections. The data set is 99.9% complete (99.2% complete in the last shell, 1.69–1.64 Å), with an average multiplicity of 4.2 and an *R*<sub>merge</sub> of 5.7%.

The high sequence homology (see below) of TMA to pig pancreatic α-amylase [22], whose crystal structure coordinates

are available, presumably will permit to solve the TMA structure employing Patterson search techniques.

#### 4. Discussion

The α-amylase from *Tenebrio molitor* larvae, has so far been one of the most studied insect α-amylases. However, all biochemical and physiological studies were performed

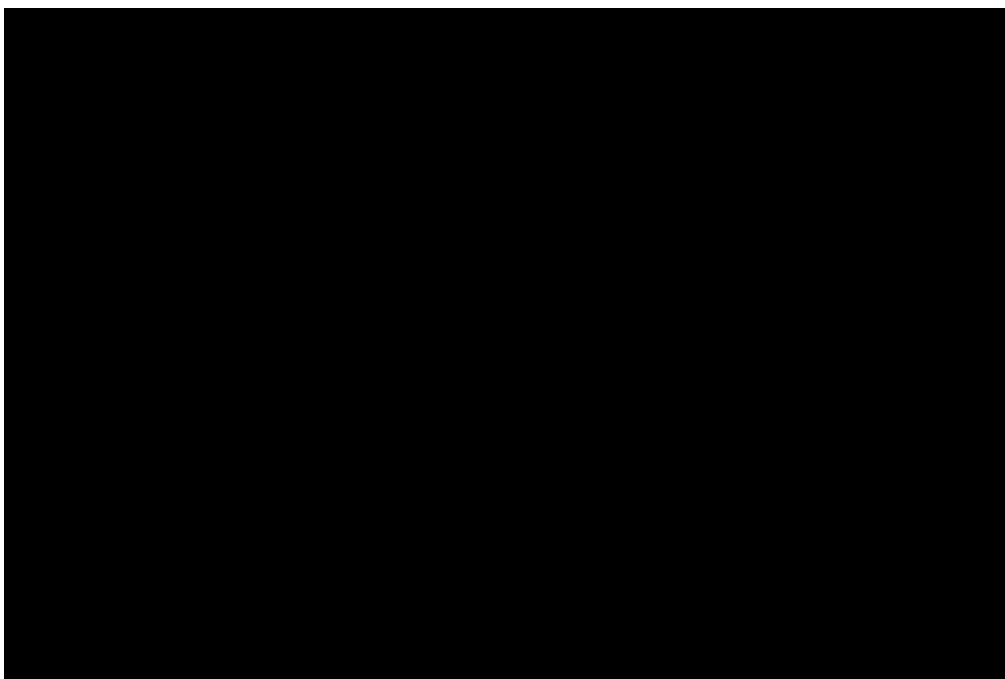


Fig. 4. Orthorhombic crystal of TMA with dimensions of approximately  $1.5 \times 0.4 \times 0.4$  mm. Conditions for crystallization are given under Section 2.

without knowledge of the enzyme's sequence and tertiary structure. In order to obtain large amounts of the homogeneous protein for sequence determination and crystallization, we established a conventional 5-step purification method (Fig. 1). Since a N-terminal protein sequence could only be obtained after treatment of TMA with pyroglutamate aminopeptidase, we deduced that a N-terminal glutamine is post-translationally modified to a pyroglutamate (5-oxo-proline). N-terminal pyroglutamates have also been reported for mammalian  $\alpha$ -amylases [18–20] and are presumably present in most insect  $\alpha$ -amylases. This feature protects the enzymes from being degraded by aminopeptidases in the digestive tract.

The complete amino acid sequence of TMA was obtained from N-terminal sequencing and translation of the cDNA sequence (Fig. 2). The comparison of the primary structure of TMA with those of other  $\alpha$ -amylases reveals that the sequence identity to the known insect  $\alpha$ -amylases is 57–79%. In the case of the mammalian enzymes, for example pig pancreatic  $\alpha$ -amylase (PPA), the sequence identity is still 54% (Fig. 3). All residues which are known from the three-dimensional structure of PPA to be involved in catalysis and binding of the single  $\text{Ca}^{2+}$  and  $\text{Cl}^-$  ions are identical in TMA and the other insect  $\alpha$ -amylases (Fig. 3). The only exception is the chloride binding residue Arg<sup>337</sup> (numbering according to PPA), which is replaced by a glutamine in the sequence of the butterfly *Ostrinia nubilalis* (amino acid numbering according to PPA). The sequence comparison also shows that all regular secondary structure elements defining the general fold of mammalian  $\alpha$ -amylases [23] are most likely conserved in the insect enzymes. The most striking difference between mammalian and insect  $\alpha$ -amylases is the presence of additional loops in the vicinity of the active site of the mammalian enzymes. Specially, the beetle  $\alpha$ -amylases lack a glycine-rich loop (residues 304–306 in PPA), which has been proposed to be involved in a 'trap-release' mechanism of substrate and product

[20]. The involvement of this segment in binding or cleavage of substrate should be reflected by the catalytic parameters of PPA in comparison to those of TMA, which is lacking this loop. At 25°C and the pH of maximal activity, pH 6.9 for PPA [24] and pH 5.8 for TMA [1], both  $\alpha$ -amylases display similar catalytic efficiency for the small substrate NPM (PPA:  $K_M = 6$  mM,  $k_{\text{cat}} = 0.06$  s<sup>-1</sup>,  $k_{\text{cat}}/K_M = 10$  M<sup>-1</sup> s<sup>-1</sup> [25]; TMA:  $K_M = 43$  mM,  $k_{\text{cat}} = 0.83$  s<sup>-1</sup>,  $k_{\text{cat}}/K_M = 19$  M<sup>-1</sup> s<sup>-1</sup>). Comparable relative activities of both  $\alpha$ -amylases were measured before for the substrate starch [1,24]. Therefore, it can be excluded that the glycine-rich loop plays an important role in binding or processing of large and small substrates.

Yellow meal worms almost exclusively live on seed products that contain high amounts of  $\alpha$ -amylase inhibitors. Therefore, it is essential that the activity of TMA, which is the most important digestive enzyme of *T. molitor*, is not completely inhibited by plant inhibitors. A detailed inspection of the X-ray structures of free PPA and PPA in complex with the microbial  $\alpha$ -amylase inhibitor tendamistat [26] and the  $\alpha$ -amylase inhibitor from the bean *Phaseolus vulgaris* [27] reveals that the glycine-rich loop and the loop at position 347–354, which is absent in all insect  $\alpha$ -amylases, and the loop 140–148, interact with these inhibitors (Fig. 3). Therefore we suggest that insect  $\alpha$ -amylases lacking these loops have a lower affinity to the proteinaceous plant inhibitors than the mammalian enzymes. The high resolution crystal structure of uncomplexed TMA, which is under way, should provide further insight into the structural differences between insect and mammalian  $\alpha$ -amylases and the influence of these differences on binding of substrate and inhibitors.

Since TMA, the main digestive enzyme of a cosmopolitan grain consuming insect, shows high homology to other insect  $\alpha$ -amylases, we believe that it constitutes a very suitable model protein for studying the interactions between insect  $\alpha$ -amylases and proteinaceous inhibitors.

**Acknowledgements:** Discussions with H. Fischer, M. Huber-Wunderlich and U. Müller are gratefully acknowledged.

## References

- [1] V. Buonocore, E. Poerio, V. Silano, M. Tomasi, *Biochem. J.* 153 (1976) 621–625.
- [2] U.S. Srivastava, P.D. Srivastava, *Beitr. Entomol.* 11 (1961) 15–20.
- [3] S.W. Applebaum, *J. Ins. Physiol.* 10 (1964) 897–906.
- [4] R. Shainkin, Y. Birk, *Biochim. Biophys. Acta* 221 (1970) 502–513.
- [5] V. Silano, F. Pocchiari, D.D. Kasarda, *Biochim. Biophys. Acta* 317 (1973) 139–148.
- [6] V. Buonocore, E. Poerio, W. Pace, T. Petrucci, V. Silano, M. Tomasi, *FEBS Lett.* 67 (1976) 202–206.
- [7] V. Silano, J.C. Zahnley, *Biochim. Biophys. Acta* 533 (1978) 181–185.
- [8] V. Buonocore, F. Grmaenzi, W. Pace, T. Petrucci, E. Poerio, V. Silano, *Biochem. J.* 187 (1980) 637–645.
- [9] F. García-Maroto, P. Carbonero, F. García-Olmedo, *Plant Mol. Biol.* 17 (1991) 1005–1011.
- [10] U.K. Laemmli, *Nature* 227 (1970) 680–685.
- [11] Frank, G., Redweik, U., Rumbeli, M., Schoedon, G. and Blau, N. (1993) in: K. Imahori and F. Sakiyama (Eds.), *Methods in protein sequence analysis*, Plenum Press, New York, pp. 79–85.
- [12] Fischer, S., Reimann, F. and Wittmann-Liebold, B. (1989) in: B. Wittmann-Liebold (Ed.), *Methods in protein sequence analysis*, Springer, Berlin, pp. 98–107.
- [13] Frank, G. (1989) in: B. Wittmann-Liebold (Ed.), *Methods in protein sequence analysis*, Springer, Berlin, pp. 116–121.
- [14] M. Wunderlich, R. Glockshuber, *J. Biol. Chem.* 268 (1993) 24547–24550.
- [15] S.C. Gill, P.H. von Hippel, *Anal. Biochem.* 182 (1989) 319–326.
- [16] Leslie, A.G.W. (1991) in: D. Moras, A.D. Podjarny, and J.C. Thierry (Eds.), *Crystallographic computing V*, Oxford University Press, Oxford, pp. 27–38.
- [17] CCP4 (1994) *Acta Crystallog. sect. D* 50, 760–763.
- [18] I. Kluh, *FEBS Lett.* 136 (1981) 231–234.
- [19] G.D. Brayer, Y. Luo, S.G. Withers, *Prot. Sci.* 4 (1995) 1730–1742.
- [20] N. Ramasubbu, V. Paloth, Y. Luo, G.D. Brayer, M.J. Levine, *Acta Cryst. Sect. D* 52 (1996) 435–446.
- [21] K. Maskos, M. Huber-Wunderlich, R. Glockshuber, *FEBS Lett.* 397 (1996) 11–16.
- [22] M. Machius, L. Vértessy, R. Huber, G. Wiegand, *J. Mol. Biol.* 260 (1996) 409–421.
- [23] S.B. Larson, A. Greenwood, D. Cascio, J. Day, A. McPherson, *J. Mol. Biol.* 235 (1994) 1560–1584.
- [24] A. Levitzki, M.L. Steer, *Eur. J. Biochem.* 41 (1974) 171–180.
- [25] Maskos, K. (1995) Ph.D. Thesis, ETH-Zürich, Switzerland.
- [26] G. Wiegand, O. Epp, R. Huber, *J. Mol. Biol.* 247 (1995) 99–110.
- [27] C. Bompard-Gilles, P. Rousseau, P. Rougé, F. Payan, *Structure* 4 (1996) 1441–1452.
- [28] L. Pasero, Y. Mazzei-Pierron, B. Abadie, Y. Chicheportiche, G. Marchis-Mouren, *Biochim. Biophys. Acta* 869 (1986) 147–157.
- [29] Abukashawa, S., Genest, Y., Hickey, D.A. (unpublished).
- [30] Popadic A., Anderson W. (unpublished).
- [31] Foster P.G., Abukashawa, S., Magoulas, C., Hickey, D.A. (unpublished).
- [32] Foster P.G., Abukashawa, S., Hickey, D.A. (unpublished).
- [33] Foster P.G., Hickey D.A. (unpublished).