The α-amylase from the yellow meal worm: complete primary structure, crystallization and preliminary X-ray analysis

Stefan Strobl^a, Franz-Xaver Gomis-Rüth^{1,b}, Klaus Maskos^{a,b}, Gerhard Frank^a, Robert Huber^b, Rudi Glockshuber^{a,*}

^aInstitut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg, CH-8093 Zürich, Switzerland ^bMax-Planck-Institut für Biochemie, D-85152 Planegg-Martinsried, Germany

Received 10 April 1997

Abstract The α -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 α -amylases and also shows high homology to the mammalian enzymes. TMA was crystallized in form of well-ordered orthorhombic crystals of space group P2₁2₁2₁ 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 α -amylase and proteinaceous plant inhibitors on the molecular level.

© 1997 Federation of European Biochemical Societies.

Key words: Tenebrio molitor; α -Amylase; Protein sequence; Crystallization

1. Introduction

 α -Amylases (α -1,4-glucan-4-glucanohydrolases, EC 3.2.1.1) constitute a family of enzymes that catalyze the hydrolysis of α -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 α -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 α -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 CaCl₂). The homogenate was centrifuged $(23\,000 \times g, 30 \text{ min}, 4^{\circ}\text{C})$ and ammonium sulfate was added to the supernatant to 90% saturation. The precipitated protein was collected by centrifugation (48000×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 α -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 *a*-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 α -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 μ g TMA were denatured by heating for 30 min at 80°C in 11 μ l of standard buffer containing 7.4 M guanidinium chloride and 90 mM dithiothreitol (DTT). After addition of 29 μ l of distilled water, 40 μ l of 200 mM sodium phosphate (pH 8.0), 20 mM EDTA, 10 mM DTT, 10% (w/v) glycerol, and 1.25 μ 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 μ 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 polyvinylidendifluoride 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

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

- NI: 5'-ATAGCGCGAGCTCARAARGAYGCNAAYTTYGC-3' CI: 5'-TTTTCAAGAAGCTTATCANAGYTTNGCRTT-
- *NACRTGDATNGC-3'* N2: 5'-CGAATTTCTAGAGGCC*TGAACCAAAACCTCAA*-
- CAC-3'
- C2: 5'-CGTCAAGGATCCAAGCTTTTTTTTTTTTT-3'

The reactions were performed with 10 μ l of the first-strand synthesis product and 2.5 U Pwo DNA polymerase (Boehringer Mannheim) in 100 μ l of 10 mM Tris-HCl (pH 8.85) at 20°C, 25 mM KCl, 5 mM ammonium sulfate, 2 mM MgSO₄, containing 0.1 mg/ml bovine serum albumin, all NTPs (0.2 mM each) and 100 pmol of the corresponding N- and C-terminal primers (30 cycles were run under the following conditions: 95°C, 60 s; 55°C, 90 s; 72°C, 120 s).

The PCR fragment amplified with primers N1 and C1 was cut with *Ecl*136II and *Hin*dIII. The fragment amplified with primers N2 and C2 was digested with *StuI* and *Hin*dIII. Both fragments were cloned into the vector pRBI-PDI [14] which had been cut with *StuI* and *Hin*dIII.

Both strands of the amplified cDNA fragments from four independent clones were sequenced by the dideoxy method and found to be identical. The sequence data has been submitted to the PIR data base PIR2: S75702).

2.4. α-Amylase activity test

During protein purification, fractions were qualitatively tested for α -amylase activity as follows: TMA samples were incubated for 5 min in 500 µl of standard buffer (see Section 2.1.) containing 0.4% (w/v) Zulkowsky starch (Merck) (~27 saccharide units), mixed with 500 µl of 1% (w/v) dinitrosalicylic acid in 0.4 M NaOH, and incubated for 5 min at 100°C. After cooling on ice, product formation was detected by the absorbance at 546 nm ($\epsilon_{546nm} = 1230 \text{ M}^{-1} \text{ cm}^{-1}$).

2.5. Protein concentrations

Protein concentrations were measured by their absorbance at 280 nm. A value of $A_{280nm,1mg/ml,1cm} = 1.641$ was determined for native TMA according to Gill and van Hippel [15].

2.6. Determination of catalytic parameters of TMA for the substrate p-nitrophenyl-α-p-maltoside (NPM)

The $K_{\rm M}$ of TMA for the substrate NPM was determined at pH 5.8, where the activity of TMA for cleavage of starch is maximal [1]. TMA activity was assayed at concentrations between 3.4 and 8.4 nM in 20 mM MES/NaOH (pH 5.8), 100 mM NaCl, 1 mM CaCl₂ at 25°C and NPM concentrations between 2 and 100 mM. Substrate hydrolysis was followed by the increase in absorbance at 405 nm ($\epsilon_{405nm,pH5.8} = 1300 \text{ M}^{-1} \text{ cm}^{-1}$).

2.7. Crystallization and X-ray analysis

Prior to crystallization, TMA was subjected to size exclusion chromatography on a Superdex 200 HighLoad 26/60 column (Pharmacia), equilbrated with 5 mM acetic acid/NaOH (pH 5.4), 0.1 mM CaCl₂, and subsequently concentrated to approximately 63 mg/ml. Crystals were grown using the hanging drop vapour diffusion method at 22°C from 5 µl droplets, consisting of equal volumes of protein solution and precipitant (200 mM sodium acetate, 100 mM BisTris-HCl (pH 6.5), 30% (w/v) PEG 8000). The droplets were equilibrated against 500 µl of precipitant solution. Crystals were directly mounted from the drop in thin-walled glass capillaries. X-ray diffraction data were recorded on an imaging plate detector (MAR Research, Hamburg, Germany) attached to a Rigaku-Denki rotating anode generator operated at 5.4 kW providing graphite-monochromatized CuK_{α} radiation. Data were processed using the MOSFLM v. 5.23 program [16] and routines from the CCP4 suite [17].

3. Results

3.1. Purification of TMA

TMA was purified to homogeneity (Fig. 1) from the crude



Fig. 1. SDS-PAGE showing the purification of TMA. Lane 1: ammonium sulfate precipitation of the crude larval extract. Lane 2: pure TMA. Lane 3: molecular mass standard.

extract of *Tenebrio molitor* larvae by a 5-step purification protocol, using ammonium sulfate precipitation, ion exchange chromatography, hydrophobic chromatography, and gel filtration (see Section 2.1). Typically, about 100 mg pure TMA were obtained from 1 kg larvae.

3.2. Complete amino acid sequence of TMA

Initial attempts to determine the N-terminal amino acid sequence of TMA by Edman degradation failed due to the presence of a N-terminal pyroglutamate, which also occurs in mammalian α -amylases [18–20]. After enzymatic cleavage with pyroglutamate aminopeptidase, residues 2-18 of TMA could unambiguously be identified by Edman sequencing (K² D³ A⁴ N⁵ F⁶ A⁷ S⁸ G⁹ R¹⁰ N¹¹ S¹² I¹³ V¹⁴ H¹⁵ L¹⁶ F¹⁷ - E¹⁸).

The primary structure of mature TMA (Fig. 2) was obtained by amplification of the cDNA with PCR in two steps. In a first PCR, the following primers were used: (i) a degenerate N-terminal primer (N1) corresponding to the N-terminal amino acid sequence of TMA (QKDANFA) and (ii) a Cterminal primer (C1) corresponding to the identical C-termini of the related insect α -amylases from Triboleum castaneum and Drosophila pseudoobscura (AIHVNAKL). The cloned PCR product included the complete cDNA encoding residues 8-463 of mature TMA. The last eight amino acid residues were verified by amplification of the 3'-coding region of the TMA cDNA sequence with primers corresponding to the internal cDNA sequence at bases 1232-1250 (N2) and to the poly-A tail of eucaryotic mRNA (C2) (Fig. 2). The resulting PCR product (229 bp) confirmed that the carboxy-terminal residues of TMA are indeed identical to those of the α -amy-

1 CAGAAGGACGCGAATTTTGCAAGTCGTAGAAATAGCATCGTGCACTTGTTCGAATGGAAATGGAATGGAATGACATC 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 GCCGACGAATGCGAGAGATTCTTGCAGCCCCAAGGATTCGGAGGAGTTCAGATCTCTCCACCTAACGAGTAC 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 48145 TTGGTGGCGGATGGCAGACCCTGGTGGGAACGGTACCAACCCGTGAGCTACATCAACACCAGGTCTGGA 216 49 L V A D G R P W W E R Y O P V S Y T T 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 120 361 CCAGCTGTACCGTATGGTTCCGGAGATTTCCACAGCCCTTGTGAAGTCAACAACTACCAAGACGCTGACAAC 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 144504 145 V R N C E L V G L R D L N Q G S D Y V R G V L T 168 505 GACTACATGAACCATATGATCGATTTGGGGGTGGCTGGATTCAGAGTGGATGCCGCCAAACACATGTCGCCT 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 P192 577 GGAGATCTGAGTGTGATCTTCTCCGGCTTGAAAAATTTGAACACCGATTACGGCTTCGCAGACGCGCCTAGA 648 193 G D L S VIFSGLKNLNTDYGFADGAR 216 649 CCCTTCATCTACCAAGAAGTTATAGATCTGGGTGGTGACGCTATCAGCAAGAACGAGTACAAGGCTTTGGT 720 240 I Y Q E V I D L G G E A I S K N E Y 217 P F T = GFG 721 TGCGTCTTGGAATTCCAGTTCGGAGTCAGTCTAGGCAACGCCTTCCAGGGTGGAAACCAGTTGAAGAATTTG 792 241 C V L E F O F G V S L G N A F O G G N O L K N L 264 793 GCGAACTGGGGTCCAGAATGGGGTCTACTCGAAGGCCTAGACGCTGTTGTGTTCGTCGTCGACAATCACGACAAT 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 CAACGTACCGGCGGGGGGGTCAAATTTTGACGTACAAGAACCCCAAGCCGTACAAAATGGCGATCGCTTTCATG 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 Μ 312 937 TTGGCCCATCCTTATGGCACCACAAGGATCATGTCCAGTTTTGACTTCACCGACAACGATCAAGGACCTCCT 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 CAAGATGGCAGCGGCAACTTGATTTCTCCTGGAATCAATGACGACAACACCTGTAGCAATGGATACGTCTGC 1080 3370 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 GAGCACCGTTGGAGGGCAGGTTTACGGAATGGTGGGATTCAGAAATGCGGTTGAAGGGACACAAGTAGAGAAAT 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 N384 1153 TEGTEGTECAATGATGACAACCAGATCECCTTCAGTCGAGGAAGTCAACGATTTGTAGCGTTTACCAACGGT 1224 385 W W S N D D N O I A F S R G S Q G F V A F T N G 408 1225 GGAGACTTGAACCAAAAACCTCAACACTGGACTTCCTGCTGGTACTTATTGCGACGTTATCTCCGGAGAGTTG 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 TCCGGTGGGTCTTGCACCGGCAAATCTGTAACAGTTGGAGATAACGGATCTGCTGATATTTCTTTGGGAAGT 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 GCCGAAGATGATGGAGTCCTAGCTATCCATGTTAACGCAAAATTGTAAATAATGTAAAAGACGATTCGAA 1440 457 A E D D G V L A I H V N A K L 471 1441 CAGT - polv 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 deoxyoligonucleotide primer which was used for amplification of the gene by PCR and which had been deduced from N-terminal Edman sequencing.

lases from Triboleum 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 α -amylase sequences from insects and exhibits the same characteristic differences from the mammalian enzymes as the other insect α -amylases (Fig. 3).

3.3. Determination of catalytic parameters of TMA

Catalytic parameters of TMA were determined with the disaccharide substrate *p*-nitrophenyl- α -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 α -amylase substrate,

but allows on-line detection and quantification of substrate hydrolysis [21]. The $K_{\rm M}$ value for the cleavage of NPM was found to be 43 ± 7 mM at 25°C and pH 5.8 and $k_{\rm cat}$ was determined to be 0.83 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 P2₁2₁2₁ with cell constants a = 51.24 Å, b = 93.46 Å, c = 96.95 Å, 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 Å³/Da (corresponding to 45% solvent content), explaining the high diffraction power up to more than 1.6 Å

	1 20			40			60			80 98			
PPA TMA TCA DPA ONA AMA CTA	QYAPQTQSGRTDIVHLFEWRWVDIALECERYI QKDANFASGRNSIVHLFEWKWNDIADECERFI QKDPHFAADRNSIVHLFEWKWDIADVCERFI QFNTNYVGGRSGMVHLFEWKWDDIADECERFI FKNPHYSGDRTTMVHLFEWKWDDIADECERFI QHDPHFVRGHSTIVHLFEWKWSDIADECERFI QFOTHQWADRCGIVHLFEWRWNDIADECERFI				LGPKGF LQPQGF LGPQGY LGPNGF LGPKGY LAPKGY	GGVQVSPPNE GGVQISPPNE GGVQISPPNE AGLQVSPVNE GGIQISPPNE GGVQLSPVNE AGVQVSPPTE	NVVVT YLVT NLVVT NAVT NLIIR NIVIRLA NAIVWSP	NPSRPWW ADGRPWW SSNRPWW KSGRPWW AHNRPWW DGSRPWW VR.RPWW	ERYQPVS ERYQPVS ERYQPVS ERYQPIS ERYQPMS ERYQPIS	SYKLCT SYIINT SYILNT SYKLTT SYRLIT SFKLDT SYELTT	RSGNENEFF RSGDESAFT RSGDETAL RSGNEQF7 RSGSEAEF7 RSGSEAEF7 RSGSEAEF7	RDMVTRC IDMTRRC ADMISRC ASMVRRC INMVRRC ADMSRRC ASMVRRC	NNVGVRIYVDAV NDAGVRIYVDAV NAVGVRIYVDIV NNAGVRTYVDVI NNVGVRIYVDAI NAAGVRLYVDII NDVGVRIYVDIV
		Αβ1		Αα1		Αβ2		A3	10 ² a			Aα+3 ₁₀ 2	b Aβ3
99 120					140		16	C		180	0	198	
PPA TMA TCA DPA ONA AMA CTA	+ INHMCGSGA INHMTGM. FNHMAAD. INHMTGTW INHMGATQI INHMAAI.	AAAGTGTTGT .NGV.GTS .GGT.GTA .GGTYGTA .SENVGTA PVEPAIGTG .TGDGGTG	GSYCNPG GSSADHI GSQADRI GSTASPS GSTATFG SSTAIPS GSTGSSQ	SNREFPAV SGMNYPAV SKNYPAV SKSFPAV SQWSYPAV SDRQFPAV 2TLSFPGV	· PYSAWI PYGSGI PYGSGI PYSSLI PYGWNI PFGWPI PFSALI	OFNDGKCKTAS OFHSP.CEV OFNSP.CEV OFNPT.CGI OFNPP.CAI OFNPP.CAI	GGIESYN NNYQ NNYQ SNYN QGSDYAN NDWG NDWG	+ DPYQVRD DADNVRN DASNVRN DANQVRN NAERVRN NAEQIRN NAIEVRN	CQLVGLI CELVGLF CELVGLF CELSGLK CELVGLF CWLVGLF = ^^^	+ JDLALE RDLNQG ADLNQG RDLNQG IDLNQG PDLNQG ^^	KDYVRSMI SDYVRGVLJ SDYVRSKII NSYVQEKVS TEHVRTMIV VPWVRDRV VQWVRDKIV	ADYLNKL I DYMNHM I EYMNHL SDFLNHL NYMNHL VDFLNHL VELFNKL	- * IDIGVAGFRUDA IDLGVAGFRVDA IDLGVAGFRVDA IDLGVAGFRVDA IDLGIAGFRIDA IELGVAGFRVDA IGMGVAGFRVDA
	вβ1 і	3β2	вβЗ	вβ4	вβ5			Βα1	вβ	36	2	Αα3	Αβ4
199			220		*	240		:	260		2	280	296
PPA TMA TCA DPA ONA AMA CTA	SKHMWPGD AKHMSPGDI AKHMWPADI AKHMWPADI AKHMWPGDI AKHMWPADI VKHMWPGDI ^ =====	IKAVLDKLH LSVIFSGLK LEAIYGSLK LGVIYGRLK LRVIYERLR LAVIFGRLN LQAIYSRLE	INLNTNW. INLNTDYG INLNTDHG INLNTDHG INLNTNHG INLNTAYG INLNTAYG	FPAGSRP SFADGARP SFLDGQKP SFASGSRP SFPAGARP SFPAGSRA SFPPNAQA	FIFQEV FIFQEV FIFQEV PYIVQEV FLAQEV FLAQEV FLAQEV	VIDLGG.EAIK VIDLGG.EAIS VIDLGG.EAIS VIDMGG.EAIS VIDLGG.EAVT VIDMGAHEVVR VIDLGG.EAVT	SGEYFSN KNEYTGF KHEYTGF KSEYTGL KHEYTPL KFEYTFL RDEYTHL =======	GRVTEFK GCVLEFQ GTVIEFQ GAVTEFR AAVTEFK GTVTEFM GTVTEFR ^^^==	YGAKLGI FGVSLGN YGLSLGN HSDSIGN FGMELSF FSHYLGF FSAEIGF =======	VVRKW JAFQ JAFQ KVFR RAFQ RAFG VFR	SGEKMSYLH GGNQLKNLJ GGNQLANLJ GKDKLTYMS RGNQLRWLM GNDALRWLS GHDRLANLH ======	KNWGEGW ANWGPEW SNWGTGW VNWGPQW SNFGEAW RNWGEAW	GFMPSDRALVFV GLLEGLDAVVFV IGLLOCLDAVAFI GFAASDRSLVFV IGLASRDAFVFV IGLLASRDAFVFV IGFLPSHLALVFV IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
	A	α+3 ₁₀ 4			Αβ5		A3 ₁₀ 5	аβ6 .	A3 ₁₀ +α6	ōa	A310	6b	Αβ7
2	297			320		_340		360			380		396
PPA TMA TCA DPA ONA AMA CTA	DNHDNQRGI DNHDNQR. DNHDNQR. DNHDNQRGI DNHDNQRGI DNHDNQRGI DNHDNQRGI ========	HGAGGSSII TGGSQII TGGSQII HGAGGADVI HGAGG.NII GNYGII HGAGGPNII	TFWDAYF TYKNPKI TYKNPKI TYKVPKQ THRQPKI TYKQPKI THQLARN	KLVAVGF PYKMAIAF PYKMAIAF PYKMASAF SYKAAIAF PYKMATAF JYKMATAF	MLAHPY MLAHPY MLAHPY MLAHPY MLAHPY AAAYPY MLAFPE	GFTRVMSSYR GTTRIMSSFD GTTRLMSSFA GTPRVMSSFS GEPQLMSSYS GQLRIMSSFA GUKRVMSSFY AA8a	WARNFVN FTD FDN FTD FTD FDD	GEDVNDW NDQ TDQ TDQ FDQ FDQ	IGPPNNN SPPQDGS SPPQDDA SPPTTDC SPPMNNN SPPSDAC SPPQDAN	IGVIKE SGNLIS AGNLIS SQNIAS JQDIIS JGNLLS JGNLIP	VTINADTT(PGINDDNT(PSINDDGT(PTFNSDKS(PSINSDGT(PIINPDKT(PVINERGL(CGNDWVC CSNGYVC CGNGYGC CGGGWVC CGNGWVC CGGGWVC CDNGYVC	EHRWRE I RNMVW EHRWRQVYGMVG EHRWRQ I FNMVG EHRWRQ I YNMVA EHRWRQ I FQMVQ EHRWRQMYSMIH EHRWRQMFNMVE
	A31074			ACTD		лроа			Aboo	Арос			Auo
3	97		420)		440			460			480	496
PPA TMA TCA DPA ONA AMA CTA	FRNVVDGE FRNAVEGT FRNAVQGT FRNAAGDA FRNVAGNT FRNLAWGT FSNVVRGT	PFANWWDNG QVENWWSNE GIENWWSDG DLQNWWSNG GLNDWWDNG QVNDWWDNG GVNDWDNG	SNQVAFC DDNQIAFS SNQQIAFC SNQISFS SNQIAFC SNNQIAF7 SANQMAFC	GRGNRGFI GRGSQGFV GRGNKGFV GRGNRAFV CRGGQAFI ARGNVGFV CRGNRGFI	VFNNDI AFTNG AFTIG AFNNDI AFNNDI AFNNEI AFNLES	WQLSSTLQTG GDLNQNLNTG YDLNQHLQTG YDLNSSLQTG WDLSQTLQTC FDMNVILQTG FFNLSQTLQTC	LPAGTYC LPAGTYC LPAGSYC LPAGTYC LPAGQYC LPAGIYC LPAGTYC	DVISGDK DVISGEL DVISGNA DVISGEK DIISGSR DVISGAR DVISGVK	VGNSCTO SGGSCTO ENGSCSO SGSSCTO SGNGCTO EGETCTO QGSTCTO	GIKVYV GKSVTV GKTITV GKVVTV GKVVTV GLQVIV GGTVTV	/SSDGKAQF /GDNGSADI /GGDGYADI /GSDGRASI /GNDGRAHI /EPNGFASI /GANGLAQI /GANGLAQI	SISNSAE SLGSAEI SLGANEE SISSSAE SVGANEY SIRANAE FIQTSAW	DPFIAIHAESKL DGVLAIHVNAKL DGVIAIHVNAKL DGVLAIHVNAKL DMMLAIHVGTQ. DGVLAIHADSRI
	Αα8	Cβ1	Сβ2	Cf	33	Cβ4	Cβ5	сβ	6 Сβ7	сβ8	в сβ9		Сβ10

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 Ca²⁺ and Cl⁻ 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 (Pir1:alpgp) [28], 53.8%/71.5%; TMA-Tenebrio molitor α -amylase (PIR2: S75702); TCA-Triboleum castaneum α -amylase (em-in:u04271_B) [29], 78.6% (87.5%; DPA-Drosphila 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_{α} radiation. A total of 244244 measurements was collected and processed to 1.64 ${\rm \AA}$ 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_{merge} 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 posttranslationally modified to a pyroglutamate (5-oxo-proline). N-terminal pyroglutamates have also been reported for mammalian α -amylases [18–20] and are presumably present in most insect α -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 α -amylases reveals that the sequence identity to the known insect α -amylases is 57–79%. In the case of the mammalian enzymes, for example pig pancreatic α -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 Ca²⁺ and Cl⁻ ions are identical in TMA and the other insect α -amylases (Fig. 3). The only exception is the chloride binding residue Arg³³⁷ (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 α -amylases [23] are most likely conserved in the insect enzymes. The most striking difference between mammalian and insect α -amylases is the presence of additional loops in the vicinity of the active site of the mammalian enzymes. Specially, the beetle α -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 α -amylases display similar catalytic efficiency for the small substrate NPM (PPA: $K_{\rm M} = 6 \text{ mM}, k_{\rm cat} = 0.06 \text{ s}^{-1}, k_{\rm cat}/K_{\rm M} = 10 \text{ M}^{-1} \text{ s}^{-1}$ [25]; TMA: $K_{\rm M} = 43 \text{ mM}, k_{\rm cat} = 0.83 \text{ s}^{-1}, k_{\rm cat}/K_{\rm M} = 19 \text{ M}^{-1} \text{ s}^{-1}$). Comparable relative activities of both α -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 α -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 Xray structures of free PPA and PPA in complex with the microbial α -amylase inhibitor tendamistat [26] and the α -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 α -amylases, and the loop 140–148, interact with these inhibitors (Fig. 3). Therefore we suggest that insect α -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 α -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 α -amylases, we believe that it constitutes a very suitable model protein for studying the interactions between insect α -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. Entemol. 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., Rümbeli, 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. Vértesy, 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, Biochym. 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).