WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

SHORT COMMUNICATION

The noncatalytic triad of α-amylases: A novel structural motif involved in conformational stability

Jean-Claude Marx,[1] Johan Poncin,[1] Jean-Pierre Simorre,[2] Pramod W. Ramteke,[3] and Georges Feller[1]*

[1] Laboratory of Biochemistry, University of Liège, Liège, Sart-Tilman, Belgium

[2] Laboratory of Nuclear Magnetic Resonance, Institute for Structural Biology Jean-Pierre Ebel, Grenoble, France

[3] Department of Biotechnology, Allahabad Agricultural Institute, Allahabad, India

ABSTRACT

Chloride-activated α-amylases contain a noncatalytic triad, independent of the glycosidic active site, perfectly mimicking the catalytic triad of serine-proteases and of other active serine hydrolytic enzymes. Mutagenesis of Glu, His, and Ser residues in various α-amylases shows that this pattern is a structural determinant of the enzyme conformation that cannot be altered without losing the intrinsic stability of the protein. ^{1}H-^{15}N NMR spectra of a bacterial α-amylase reveal proton signals that are identical with the NMR signature of catalytic triads and especially a deshielded proton involving a protonated histidine and displaying properties similar to that of a low barrier hydrogen bond. It is proposed that the H-bond between His and Glu of the noncatalytic triad is an unusually strong interaction, responsible for the observed NMR signal and for the weak stability of the triad mutants. Furthermore, a stringent template-based search of the Protein Data Bank demonstrated that this motif is not restricted to α-amylases, but is also found in 80 structures from 33 different proteins, amongst which SH2 domain-containing proteins are the best representatives.

Proteins 2008; 70:320–328.

Key words: α-amylase; protease; catalytic triad; protein stability; low barrier hydrogen bond.

INTRODUCTION

The catalytic triad forming the active center of serine proteases has been known for decades and the elucidation of its mechanism was one of the main achievements of modern enzymology.[1,2] In these enzymes, a charge-relay network among Asp, His, and Ser residues increases the nucleophilicity of the reactive serine side chain in the first step of peptide bond hydrolysis. A similar, catalytic and serine-based triad is found in the active site of some other hydrolytic enzymes such as cholinesterases,[3] cutinases,[4] or lipases.[5,6] Much more surprising was the finding of a constellation of Glu, His, and Ser residues in starch-degrading α-amylases,[7] perfectly adopting the spatial arrangement of the catalytic triad and having the residue composition of some lipases[8] or acetylcholinesterases,[9] that is, Glu instead of Asp present in proteases [Fig. 1(A) and S1]. This constellation is located at 22 Å from the α-amylase active site, close to the surface of the $(\beta/\alpha)_8$ central barrel, but has a low solvent accessibility (Fig. S2). The main difference with the catalytic triad is found at the level of the His residue, which adopts a flipped

Abbreviations: AHA, α-amylase from *Pseudoalteromonas haloplanktis*; AmyD, α-amylase from *Drosophila melanogaster*; Et-G7-pNP, 4-nitrophenyl-α-D-maltoheptaoside-4,6-*O*-2-ethylidene; LBHB, low barrier hydrogen bond.

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat

Grant sponsors: Fonds National de la Recherche Scientifique, Belgium; Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche, Grand-Duchy of Luxembourg; Grant sponsor: Access to Research Infrastructures activity in the 6th Framework Program of the EC; Grant number: RII3-026145, EU-NMR.

*Correspondence to: Dr. Georges Feller; Laboratory of Biochemistry, Institute of Chemistry B6a, B-4000 Liège-Sart Tilman, Belgium. E-mail: gfeller@ulg.ac.be

Received 18 January 2007; Revised 15 March 2007; Accepted 2 April 2007

Published online 29 August 2007 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/prot.21594

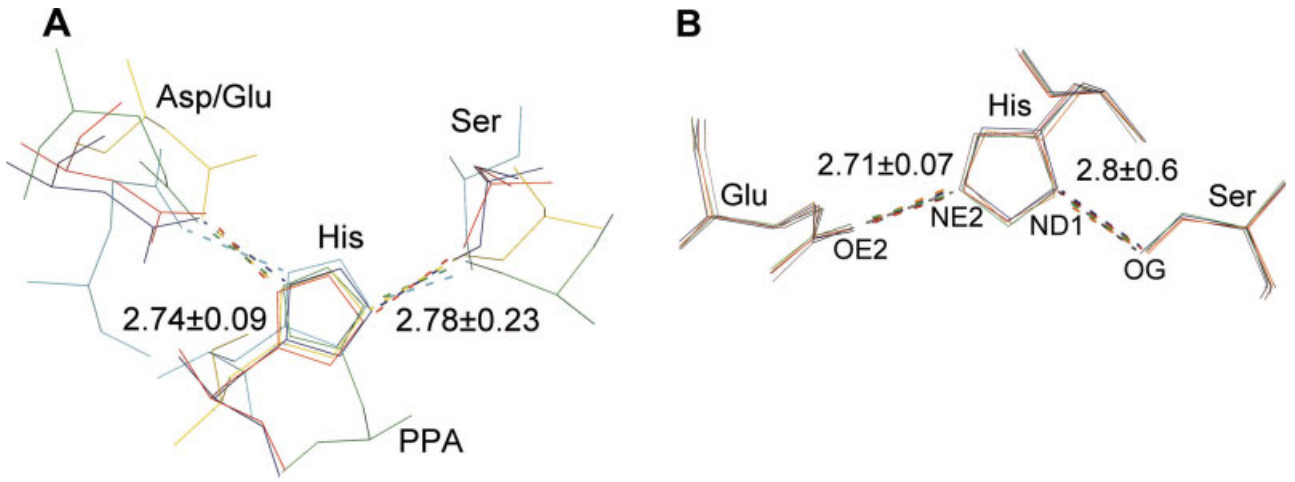

Figure 1

*Structural superimposition of catalytic and noncatalytic triads. **A**, the noncatalytic triad of pig pancreatic α-amylase (PPA, black, labeled) is superimposed with the active site catalytic triads of trypsin (blue), chymotrypsin (orange), subtilisin (green), and lipase (red). **B**, superimposition of the noncatalytic triads in the five crystal structures of chloride-dependent α-amylases: porcine pancreatic (orange), human pancreatic (green), human salivary (blue), insect (red), and bacterial (black) α-amylases. Interatomic distances are given in Å and the four functional atoms are indicated in panel B.*

position with His ND1 facing Ser instead of facing the carboxylate in the charge relay system, as shown in Figure 1(A) and S1. Although slightly differing by the side chain orientation, these triads (catalytic or not) are characterized by the perfect conservation of the geometry and interatomic distances among the four functional atoms, that is, the carboxylate oxygen of Asp (OD2) or Glu (OE2), the imidazole nitrogen atoms (ND1, NE2) of His, and the hydroxyl oxygen (OG) of Ser. Noncatalytic triads have been previously reported in proteins[10,11] but all these motifs were strongly distorted and not perfectly superimposable with the catalytic triads, in contrast with the motif found in chloride-activated α-amylases.

α-Amylases (α-1,4-glucan-4glucanohydrolases, EC 3.2.1.1) are widely distributed in microorganisms, plants, and animals. They catalyze the hydrolysis of internal α(1,4)-glycosidic bonds with net retention of the anomeric configuration in starch and related polysaccharides.[12–14] These enzymes are typically monomers of about 50 kDa exhibiting a central domain A formed by a $(\beta/\alpha)_8$ barrel, a small β-pleated domain B protruding between β3 and α3 delineating the active site, and a C-terminal globular domain C consisting of a Greek key motif[15] (Fig. S2). These enzymes also require at least one tightly bound calcium ion for structural integrity. Amongst these enzymes of various origin, one group of α-amylases also requires a chloride ion for enzymatic activity, that acts as an allosteric activator.[16,17] This unusual bound anion was observed at close proximity of the active site in the crystal structures of porcine pancreatic,[18] human pancreatic,[19] human salivary,[20] insect,[21] and bacterial[7,22] α-amylases. Site directed mutagenesis of the chloride protein ligands combined with biochemical and crystallographic studies have shown that this anion raises the pK_a of the catalytic Asp proton donor to higher values, allowing optimal activity near neutrality, but also shields this side chain from unfavorable electrostatic interactions.[23–26] A bioinformatic search based on the occurrence of the conserved chloride ligands within the α-amylase sequences has revealed that the chloride-dependent α-amylase family is composed of all animal α-amylases and of some bacterial enzymes[27] such as the α-amylase from *Pseudoalteromonas haloplanktis* (AHA) used in this study. In the latter case, it has been suggested that the occurrence of an animal-type α-amylase in bacteria originates from horizontal gene transfer.[28]

Besides these structural and phylogenetic relationships, the chloride-dependent α-amylases are also characterized by the strict and unique conservation of the aforementioned noncatalytic triad in the five high-resolution crystal structures solved so far[7,18–22] [Fig. 1(B) and S3]. In addition, more than 70 amino acid sequences of chloride-dependent α-amylases are known and the three residues forming the triad are found to be strictly conserved,[27] revealing that the occurrence of this constellation is not fortuitous. Since the first mention of noncatalytic triads in proteins,[10,11] their possible significance (fortuitous or functional) has remained obscure. The discovery of the α-amylase model of noncatalytic triads prompted us to investigate the function and occurrence of such an unusual motif.

MATERIALS AND METHODS

Mutagenesis and enzyme expression

Site directed mutagenesis of the *P. haloplanktis* α-amylase gene in the vector pαH12wt*, enzyme expression in *E. coli* RR1 cells at 18°C and α-amylase purification were performed as described.[29] Western blots of culture samples were performed using rabbit anti-AHA IgG as primary antibodies and alkaline phosphatase conjugated anti-rabbit IgG as secondary antibodies. Random mutagenesis of pαH12wt* was carried out in Epicurian Coli XL1-Red competent cells (Stratagene) in the conditions recommended by the supplier. The pool of randomly mutated plasmids was transformed in *E. coli* RR1 and the transformants were subcultured at 37°C for 48 h in 96-well uncoated micro-plates containing 200 µL of Luria–Bertani medium, with the wild-type and the Ser303Ala mutant plasmids as controls. The Et-G7-pNP reaction mixture (see later) was added directly to the microcultures and absorbance variation at 405 nm was recorded after 15 min. Under these conditions, cells expressing the wild-type α-amylase are active whereas those expressing the Ser303Ala mutant are inactive. The gene *amy-p* coding for *Drosophila melanogaster* α-amylase was kindly provided by Dr. Jean-Luc Da Lage, CNRS, Gif-sur-Yvette, France. The sequence of native signal peptide was replaced by the sequence of the bacterial *P. haloplanktis* α-amylase signal peptide and the gene was transferred into a pET-22b vector for induced production in *E. coli* BL21(DE3) cells. Mutations of the triad were introduced with the Quikchange Mutagenesis kit from Stratagene.

Enzyme assays

Spectrophotometric determination of α-amylase activity and of the kinetic parameters were carried out using the Et-G7-pNP Amyl kit (Roche) as described.[23] Substrate specificity for polysaccharides was determined by isothermal titration calorimetry,[30] using a MCS microcalorimeter (Microcal) and soluble starch (2%) or oligosaccharides (5 m*M*) as substrates. Agar-immobilized starch, casein, and emulsified tributyrin were used to detect traces of amylolytic, proteolytic, and lipolytic activities, respectively. The substrates in 50 m*M* HEPES, pH 7.5 were autoclaved with 1.5% agar, and poured into Petri dishes. Filter-sterilized (0.22 µm) protein samples were loaded onto the solid medium and incubated at 25°C up to 1 month. Activity appeared as halos of hydrolysis. Amylolytic activity of recombinant *E. coli* colonies was detected by the same procedure on Luria–Bertani plates.

NMR spectroscopy

Purified AHA[29] was prepared in 200 µ*M* MES, 1 µ*M* $CaCl_2$, pH 5.8, and concentrated to 0.3 m*M*. Thirty microliter of D_2O were added to 430 µL of sample before running the experiment. Data were recorded at 20°C on a Varian 800 MHz Inova NMR spectrometer at the RALF-NMR facility (Grenoble). To produce ^{15}N-labeled AHA, the gene coding for AHA and its signal peptide have been inserted into the expression vector pET-22b following the engineering of an *NdeI* restriction site at the start codon. The ^{15}N-labeled enzyme was produced in M9 minimal medium containing $^{15}NH_4Cl$ as sole nitrogen source. The production yield in *E. coli* BL21(DE3) cells was ~3.5 mg/L. ^{15}N-labeled AHA was purified to homogeneity as described for unlabeled AHA.[29] Analysis of isotopic labeling was performed in a quadrupole-time of flight system (Q-TOF Ultima) after ionization by electrospray and showed a mass difference of 587 ± 4 Da between the labeled and unlabeled enzyme, indicating an isotopic labeling of 97.7 ± 0.7%.

Bioinformatics

The PDB[31] scan was performed with the constraint-based Jess algorithm,[32] a modified version of TESS.[33] Jess was kindly provided by Jonathan Barker (EBI, Hinxton, Cambridge, UK). The template consisted of the atomic coordinates of ND1 and NE2 of His337 imidazole ring and the oxygen atoms OG of Ser303 and OE2 of Glu19, based on PDB no. 1AQH of AHA. The tolerance factor Δ (the distances between pairs of atoms are within Δ Å of the corresponding distance in the template) was set to 0.5 Å. Glu OE2 atom was allowed to be replaced by any noncarbon, side chain atom without residue matching. A strict conservation was imposed on Ser OG. His ND1 and NE2 were allowed to match either with triads in protease- or in α-amylase-like orientation. Coplanarity of the heteroatoms of matching structures was checked visually with SwissPdb-Viewer. The Catalytic Site Atlas (CSA)[34] was consulted for separating catalytic from noncatalytic triads. The primary literature was consulted for crosschecking of CSA results and for the structures lacking in the CSA.

RESULTS

Mutational study of the noncatalytic triad

Both the wild-type AHA and the recombinant α-amylase produced in *E. coli* are devoid of proteolytic and lipolytic activity on agar-immobilized casein and emulsified tributyrin, which allows the detection of trace activity not recorded by usual spectrophotometric methods. This observation rules out a possible accessory hydrolytic activity in this enzyme, as already suggested by the burial and low solvent accessibility of the triad. To evaluate the function of the noncatalytic triad, the mutations Glu19Gln, Ser303Ala, and His337Asn were separately introduced in the α-amylase gene, as well as the three

DOI 10.1002/prot

combinations of double mutations (Glu19Gln/Ser303Ala, Glu19Gln/His337Asn, and Ser303Ala/His337Asn) and the triple mutation (Glu19Gln/Ser303Ala/His337Asn). These mutations eliminate the respective chemical function of the original residues while maintaining a similar steric hindrance. In the expression system used,[29] the wild-type AHA is produced at a high yield of about 200 mg/L culture. The mutants are also readily expressed in *E. coli*, as demonstrated by Western blotting using anti-AHA IgG after SDS-PAGE, and in an active form as shown by halos of hydrolysis on starch-containing LB plates. However, they display a strongly reduced stability, with a half-life of activity in cell-free extracts of one or two days, whereas the wild-type α-amylase is stable for weeks in these conditions. This weak stability is also demonstrated by immunoblotting, as these mutants become undetectable after 2–3 days of culture. The low activity of these mutants in cell-free extracts (less than 0.5% of the wild-type α-amylase expressed in the same conditions, Et-G7-pNP as substrate) also suggests that mutations in the noncatalytic triad promote strong structural alterations. The least stable and active mutants (only detectable on agar-immobilized starch) are those bearing the Glu19Gln mutation, either alone or in combination. In attempts to produce a stable mutant containing an altered triad, the following mutations were also engineered in *P. haloplanktis* α-amylase: His337, the central residue of the triad, has been subjected to saturation mutagenesis and replaced by the 19 other amino acid residues. Furthermore, two double mutants, Glu19Cys/His337Cys and Ser303Cys/His337Cys, were designed to introduce a covalent disulfide bridge within the triad. Finally, the three basic mutations Glu19Gln, Ser303Ala, and His337Asn were introduced in Mut5, a strongly stabilized variant of *P. haloplanktis* α-amylase[35] and in the closely related insect α-amylase from *Drosophila melanogaster* (AmyD) that belongs to the most stable chloride-dependent α-amylases according to microcalorimetric studies.[36] These mutants can be expressed in *E. coli* in an active form but all display the same pronounced instability and weak residual activity as described earlier. As a result, attempts to purify these mutants were unsuccessful, even using fast chromatographic techniques. These results on the 34 engineered mutants demonstrate that any alteration in the noncatalytic triad leads to a drastic destabilization of the protein structure in these α-amylases.

Mutant characterization

The mutant of the triad Ser303Ala is slightly more stable as far as the half-life of activity in cell-free extracts is concerned. The fast loss of activity of this mutant cannot be avoided by inhibition of serine-, thio-, metallo-, or acid proteases, and the stability of the enzyme in cell-free extracts cannot be improved by usual protein stabilizers such as $MgSO_4$, Na_2SO_4, sucrose, threalose, glycerol, alanine, polyethylene glycol, or acarbose, a stabilizing inhibitor of glycosidases. Trimethylamine *N*-oxide, a protecting osmolyte that has a strong ability to force proteins to fold[37] is also ineffective in stabilizing the Ser303Ala mutant. By contrast, the concentration-independent kinetic parameter K_m for the synthetic substrate Et-G7-pNP (K_m = 150 μ*M*) is identical to that of the wild-type α-amylase. Furthermore, isothermal titration calorimetry, recording the heat release of hydrolysis,[30] was used for the determination of the relative substrate specificity. The activity of the mutant Ser303Ala on the macromolecular substrate starch was set to 100% activity and the activity on shorter oligosaccharides was recorded (maltooligosaccharides from corn syrup, 64%; maltopentaose, 69%; Et-G7-pNP, 113%). These relative activities are unchanged with respect to the wild-type enzyme. Maintenance of the kinetic parameters suggests that mutants of the noncatalytic triad are produced in a native state followed by a fast and irreversible unfolding, leading to proteolysis of the unfolded state.

Random mutagenesis

In an attempt to understand the destabilizing effect of these mutations, the mutant Ser303Ala was subjected to extensive random mutagenesis in the mutator strain Epicurian Coli XL1-Red (Stratagene) with a programmed mutation frequency of one to six basepair substitutions for the AHA gene. The objective of this experiment was to obtain a secondary mutation restoring the stability of the native α-amylase or at least improving the stability of the mutant Ser303Ala. However, none of the 12,500 screened mutants displayed an improved stability compared with the Ser303Ala mutant. In control experiments on the wild-type α-amylase gene randomly mutated in the same conditions, it was found that at a mutation frequency of 1 bp substitution per gene, about 2% of the mutants possessed a significantly improved stability. The sequence of some of these mutants also confirmed the programmed mutation frequency. These results suggest that mutational perturbations in the noncatalytic triad of the α-amylase cannot be balanced by simple mutations or weak interactions in the enzyme structure.

NMR spectroscopy

At this stage, and considering the previous results, we have assumed that the noncatalytic triad contains at least one nonconventional H-bond, contributing significantly to the overall global stability of α-amylases. This hypothesis was based on the description of a strong or low barrier hydrogen bond (LBHB) in the catalytic triad of serine proteases, leading to a typical proton signal of the NMR spectrum in the 15–20 ppm range.[38] The algorithm ShiftX, allowing the calculation of chemical shifts on the basis of atomic coordinates of a molecule,[39] did

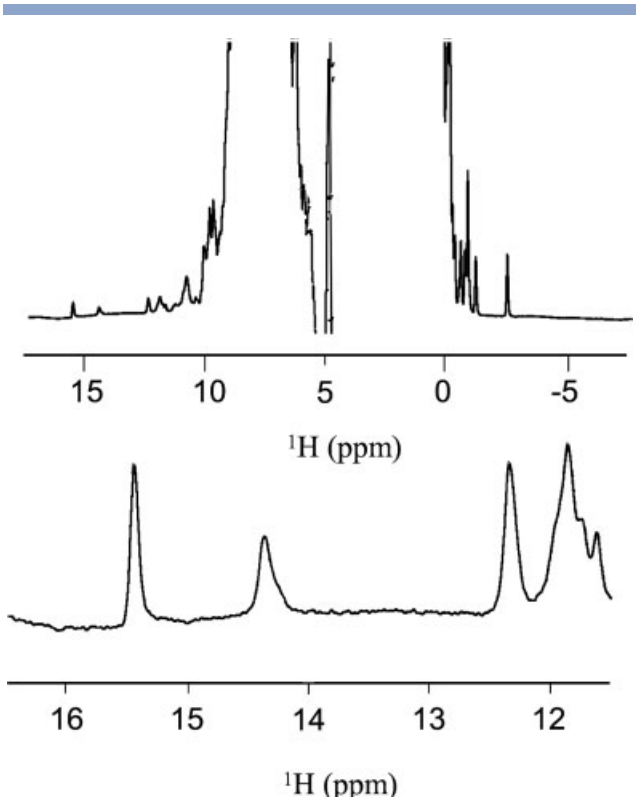

Figure 2

*800 MHz proton NMR spectra of AHA in ^1H_2O. **A**, overview of the proton signals for AHA. **B**, enlargement of the downfield spectrum in A showing two typical proton signals at 14.4 and 15.5 ppm.*

not predict any proton chemical shift higher than 10 ppm for AHA (PDB no. 1AQH). By contrast, and despite the large size of the native α-amylase (∼50 kDa) and its solubility limit (pH ∼ 5.5), the ^{1}H NMR spectra display an unusual low field proton signal (Fig. 2). This proton appears well downfield from other protons in the α-amylase, exhibiting a chemical shift δ_H of 15.5 ppm at pH 5.8. This $\delta_H$value in the α-amylase matches the chemical shift of the LBHB linking Asp and His residues in the catalytic triad of serine proteases.[40] Interestingly, the ^{1}H NMR spectrum of AHA also displays a signal at 14.4 ppm (Fig. 2), exactly as in the spectra of some serine proteases. In these cases, this signal has unambiguously been attributed to the Ser-OH proton of the catalytic triad.[41,42]

These observations prompted us to produce ^{15}N-labeled AHA to characterize the low field proton. As shown in Figure 3, the correlation spectrum of ^{15}N-labeled AHA displays a signal at 15.5 ppm (^{1}H) and 183.2 ppm (^{15}N), hence demonstrating that the proton of interest is bound to a nitrogen atom. Comparison of these experimental chemical shifts with the data of the Biological Magnetic Resonance Bank[43] indicates that the only possible N—H configuration accounting for both chemical shifts is a protonated histidine imidazole ring. In conclusion, this bacterial α-amylase contains a protonated His residue engaged in a strong H-bond, the latter displaying properties close to those of a LBHB. Unfortunately, the NMR signals could not be unambiguously attributed to the His residue composing the noncatalytic triad for several reasons: (i) two-dimensional NMR techniques failed to identify this His residue because of the size of the protein and the resulting high number of overlaying signals, and (ii) the proton NMR spectra of the triad His mutants of AHA and of AmyD should be devoid of the 15.5 ppm signal but these mutants cannot be produced (see earlier).

Noncatalytic triads in other proteins

To check the occurrence of protein structures different from α-amylases but bearing a similar Ser-His-Glu/Asp triad, the entire PDB was scanned with a stringent constraint-based algorithm. The atomic coordinates of the

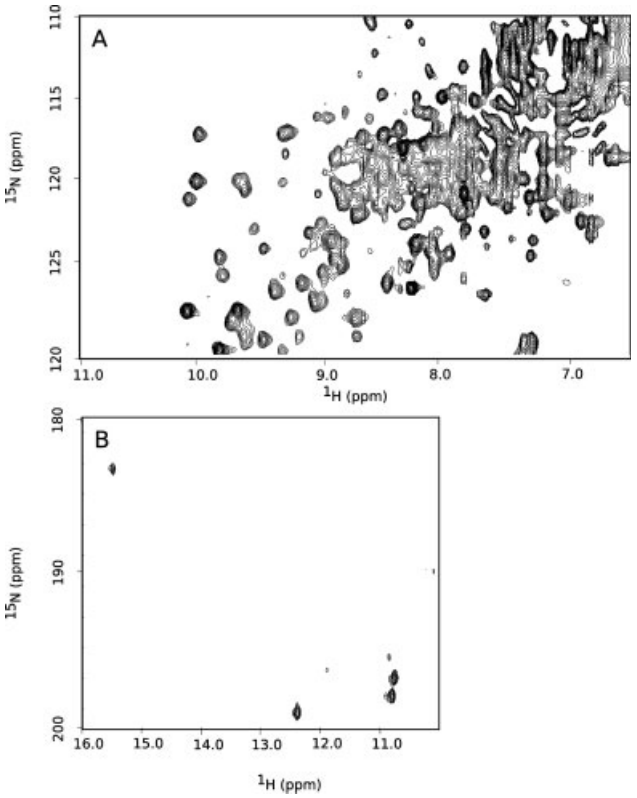

Figure 3

*800 MHz ^{15}N-^{1}H HSQC correlation spectra of ^{15}N-labeled AHA. **A**, spectrum recorded with the carrier centered on the amide proton at 120 ppm. **B**, spectrum recorded with the carrier centered at 190 ppm on the histidine ND1 chemical shift. The proton signal at 15.5 ppm (Fig. 2) is correlated with a nitrogen atom at 183.2 ppm.*

DOI 10.1002/prot

Table I
Noncatalytic Triads in Proteins

PDB no.	rmsd (Å)	Δ (Å)	Structure name as in PDB file	S	H	E/D	No. of redundancies	Redundant PDB no.
Noncatalytic amylase-like triads								
1R1S	0.08	0.05	SH2 domain	S97	H106	E71	30	1A07, 1A08, 1A1B, 1A1C, 1A1E, 1A81, 1A0T, 1A0U, 1LKK, 1LKL, 1O4G, 1O4I, 1O4L, 1O4N, 1O42, 1O46, 1OPK, 1OPL, 1CWD, 1P13, 1FYR, 1JYR, 1ZFP, 1H9O, 1QAD, 1R1P, 1R1Q, 1SPS, 2SHP
1VH4	0.13	0.14	Protein binding protein (stabilizer)	S243	H215	E187	1	
2HMY	0.16	0.35	Methyltransferase	S294	H292	E47	1	
1ODO	0.16	0.18	Cytochrome P450	S318	H291	E20	1	
1NKG	0.18	0.27	Rhamnogalacturonan lyase	S367	H365	E136	1	
1VFR	0.19	0.12	NADPH oxidoreductase	S81	H82	E54	1	
1N2K	0.19	0.38	Sulfatase	S250	H227	E253	2	1E2S
1CF2	0.22	0.18	Glyceraldehyde-3-*P* dehydrogenase	S302	H299	E301	1	
1JY1	0.24	0.2	Tyrosyl-DNA phosphodiesterase	S419	H372	E415	2	1QZQ
1IZN	0.31	0.2	Actin capping protein	S171	H152	E221	1	
1QHM	0.13	0.17	Pyruvate formate lyase	S1083	H1084	D1074	6	1CM5, 1H16, 1H17, 1H18, 2PFL, 3PFL
1BGL	0.15	0.28	β-galactosidase	S971	H972	D591	2	1JYW
1KW0	0.17	0.31	Human Phe-hydroxylase	S273	H271	D151	2	1MMT
1H46	0.18	0.28	Exoglucanase	S43	H42	D74	1	
1UKP	0.18	0.1	β-amylase	S464	H335	D275	3	1BFN, 1BYB
1ASG	0.18	0.14	Aspartate aminotransferase	S141	H145	D223	7	1ASL, 1ASM, 1CQ7, 8AAT, 1AMR, 1ARG
1DE6	0.2	0.2	Rhamnose isomerase	S177	H229	D231	2	1D8W
1SI0	0.21	0.15	Iron binding protein	S278	H267	D205	2	1PMB
1MRJ	0.23	0.18	Ribosome inactivating protein	S61	H51	D1	2	1BRY
1DEL	0.28	0.17	Deoxynucleoside-P kinase	S7	H194	D218	1	
Noncatalytic protease-like triads								
1IHU	0.07	0.13	Arsenite-translocating ATPase	S420	H148	E416	3	1F48, 1II9
2BSP	0.11	0.1	Pectate lyase	S18	H193	D150	2	1BN8
1QNQ	0.12	0.16	β-mannanase	S158	H198	D195	3	1QNO, 1QNR
1U43	0.19	0.33	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	S35	H42	D8	1	
1R8W	0.26	0.18	Glycerol dehydratase	S7	H368	D312	2	1R9D

The first four columns provide the best matching structure with its rmsd and delta values, as well as its name as it appears in the PDB file. The following three columns present the identified residues constituting the triad; the second last column shows the number of redundancies, with the PDB identifiers detailed in the last column.

four functional atoms of the triad in AHA were used as a template for this search that identified 450 matching protein structures. After the removal of enzymes bearing a catalytic triad and of all the Cl^--dependent α-amylases, 80 PDB codes were retained that fulfill the coplanarity and the distance requirements for the correct establishment of hydrogen bonds in the triad. Among these hits, 69 PDB entries, representing 28 nonredundant structures, bear an α-amylase-like noncatalytic triad whereas 11 PDB entries, representing 5 nonredundant structures, bear a protease-like noncatalytic triad (Table I). The mean distances for the structures described in Table I are 2.76 ± 0.12 Å between the His imidazole ring nitrogen atom and the carboxylic oxygen atom, and 2.85 ± 0.15 Å between the second His nitrogen heteroatom and Ser OG. The primary literature of all the retained structures never describes any resemblance with the triad, in the rare cases where the role of some of the identified residues is discussed. Without exception, these roles are described to be of a structural nature: coordination of ions, cofactor binding, stabilization through hydrogen bonding, intra- and intersubunit stabilization. In addition, several structures amongst homologous proteins are found bearing the triad (Table I) indicating that this constellation is also conserved in other proteins. The best examples are the 30 structures of SH2 domain-containing proteins, representing nine different proteins that also show the best geometry statistics for superimposition with the α-amylase template. Unfortunately, none of the structure reported in Table I has been determined by NMR, therefore impairing the search for an unusual low field proton signal in these proteins. It should also be noted that the noncatalytic triad is a unique feature of the chloride-dependent α-amylase family as no other α-amylase type from microorganisms or plants was identified by the search.

DISCUSSION

The strict conservation, in both the amino acid sequences and the crystal structures, of the 3 residues

constituting the noncatalytic triad of chloride-dependent α-amylases indicates that the occurrence of this motif is not fortuitous. The results of our site-directed and random mutagenesis studies clearly show that the noncatalytic triad fulfills an important structural role, as Glu, His, and Ser residues in various α-amylases cannot be altered without losing the intrinsic stability of the protein. Furthermore, the kinetic parameters of a transiently stable mutant of the triad indicate that this enzyme is synthesized in a native conformation, followed by a fast and irreversible unfolding. At this stage, one should wonder how a constellation of three residues can constitute a key determinant of the conformational stability in a large monomeric protein such as AHA (453 residues, 49,340 Da). This α-amylase tolerates various amino acid substitutions: more than 20 mutations have been previously introduced during various studies with some being at close proximity of the noncatalytic triad,[29,35] but none had the drastic destabilization effect displayed by mutations implicating the residues of the triad. Moreover, the two putative H-bonds linking the side chains in the noncatalytic triad are a priori unlikely to contribute significantly to the α-amylase conformational stability when compared to the 400 H-bonds identified in the crystal structure.[22] However, this effect can be explained if at least one hydrogen bond in the triad is not a conventional interaction but a strong H-bond, such as the LBHB found in the catalytic triad of serine proteases.[40]

Hydrogen bonds have been classified in three types according to their stabilization energy[38]: conventional (10–50 kJ/mol), strong or low barrier (50–100 kJ/mol), and very strong or single-well (>100 kJ/mol) hydrogen bonds, the latter providing bond strength close to covalent bonds, such as in hydrogen fluoride, but have not been reported in proteins. The most unambiguous parameter for characterizing a LBHB in proteins is the NMR chemical shift δ_H of the participating proton which ranges from 15 to 20 ppm.[38] Surprisingly, the NMR spectra of AHA revealed many features that are identical with the NMR signature of catalytic triads. First, the 15.5 ppm proton signal is typical for an LBHB between His and an adjacent carboxylate group. In the case of AHA, this proton is correlated with a nitrogen atom and is assigned to a protonated histidine. Second, a signal at 14.4 ppm is observed in the ^{1}H NMR spectrum of AHA, exactly as in the spectra of some serine proteases in which this signal has been attributed to the Ser-OH proton.

Although these NMR signals in AHA cannot unambiguously be assigned to the H-bonding network in the noncatalytic triad of the α-amylase, several observations support the occurrence of a strong H-bond between His and Glu of the triad: (i) all LBHBs detected thus far in enzymes involve a carboxyl group,[44] represented by Glu in the noncatalytic triad, (ii) analysis of the correlation between the ^{1}H chemical shift and the H-bond length for strong hydrogen bonds implicating an imidazolium ring and a carboxyl group[45] indicate that a chemical shift of 15.5 ppm corresponds to a bond length of 2.66 Å: this is in agreement with the 2.71 ± 0.07 Å measured for the triads of the five chloride-dependent α-amylase structures, (iii) The two histidine residues found in the active site of α-amylases and involved in catalysis have no carboxyl group at H-bonding distances: this rules out a possible catalytic LBHB in the active site of these enzymes, (iv) Amongst the 11 additional His (of 12 in total) in the AHA structure, three have a carboxyl group at H-bonding distance but these are nonconserved His-Glu pairs of residues within the Cl^--dependent α-amylase family.

The occurrence of a strong hydrogen bond in the noncatalytic triad can therefore account for the instability of mutants of the participating residues. Indeed, the residues in the noncatalytic triad in α-amylases bridge the N- and C-extremities of the central $(\beta/\alpha)_8$ barrel of the enzymes: removal of this interaction may lead to the opening of this major domain and to the subsequent unfolding of the structure. The functional role of a strong hydrogen bond in the noncatalytic triad in chloride-dependent α-amylases can be proposed in the light of data available for the well characterized serine proteases. On the one hand, the spatial arrangement of Glu, His, and Ser residues in the noncatalytic triad seems to provide the required environment in order to favor strong H-bonding between His and Glu. Ser is proposed to position the His side chain correctly in terms of distance and coplanarity. On the other hand, the flipped position of His in the noncatalytic triad possibly precludes abstraction of the proton from the Ser-OH group, therefore avoiding a reactivity increase of the serine side chain.

Interestingly, the geometry adopted by the three residues in the noncatalytic triad is not restricted to Cl^--dependent α-amylases but has also been detected in 80 other protein structures, accounting for 33 different proteins in the Protein Data Bank. In all these structures, the mean distance between the carboxyl group and the imidazole ring is short and compatible with the formation of a strong H-bond, whereas the mean distance between the imidazole and the Ser hydroxyl group is slightly longer, like in the α-amylase triad. This constellation is frequently found in series of homologous proteins (Table I), indicating that the triad geometry is also strictly conserved.

Although in most instances the residues forming the triad are not mentioned in the related publications, some peculiar cases are worth commenting. SH2 domain-containing proteins represent the largest number of matching structures from the search and also display the best superimposition statistics (Table I). SH2 domains contain about 100 amino acid residues and are involved in many signal transduction processes mediated by binding to phosphotyrosine-containing proteins.[46] Two interesting aspects should be mentioned: (i) the sequence position

 DOI 10.1002/prot

of His from the noncatalytic triad is variable but not its position in the 3D structures of SH2 domains, (ii) the histidine of the triad is the critical residue involved in the binding of phosphotyrosine in the majority of proteins with SH2 domains. Accordingly, the noncatalytic triad seems to be directly involved in the essential cell signaling function of SH2 domains. In the case of D-glyceraldehyde-3-phosphate dehydrogenases (GAPDH), the identified Glu residue is strictly conserved within the whole family of GAPDHs and is known to contribute to the protein stability through hydrogen bonding.[47] In addition, His and Ser residues of the triad are only conserved in hyperthermophilic archaeal GAPDHs. This suggests that the noncatalytic triad could function as a stabilizing motif in these hyperstable proteins. In the heterodimeric actin-capping protein, the identified Glu residue has been reported to be crucial for maintaining the overall architecture of one subunit and for stabilizing the dimer, as it also interacts with the second subunit. The present finding that this residue belongs to a noncatalytic triad, with His and Ser partners located on different secondary structures, suggests that this constellation could be implicated in the protein stability, as in α-amylases. In conclusion, these noncatalytic triads could represent a still not described structural motif involved in the conformational stability of proteins by enhancing the strength of interactions between the partners or with other partners located within the structure or on a different polypeptide.

ACKNOWLEDGMENTS

The facilities offered by the Institut Polaire Paul-Emile Victor, Brest, France, are acknowledged. The authors thank Dr. Jean-Luc Da Lage for providing the gene of AmyD, Dr. J. Barker for providing the algorithm Jess, and J. Torrance for providing a copy of the CSA. N. Gerardin and A. Dernier are also acknowledged for their skillful technical assistance.

REFERENCES

1. Wright CS, Alden RA, Kraut J. Structure of subtilisin BPN' at 2.5 angstrom resolution. Nature 1969;221:235–242.
2. Blow DM, Birktoft JJ, Hartley BS. Role of a buried acid group in the mechanism of action of chymotrypsin. Nature 1969;221:337–340.
3. Millard CB, Kryger G, Ordentlich A, Greenblatt HM, Harel M, Raves ML, Segall Y, Barak D, Shafferman A, Silman I, Sussman JL. Crystal structures of aged phosphonylated acetylcholinesterase: nerve agent reaction products at the atomic level. Biochemistry 1999;38:7032–7039.
4. Martinez C, De Geus P, Lauwereys M, Matthyssens G, Cambillau C. *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. Nature 1992;356:615–618.
5. Winkler FK, D'Arcy A, Hunziker W. Structure of human pancreatic lipase. Nature 1990;343:771–774.
6. Brady L, Brzozowski AM, Derewenda ZS, Dodson E, Dodson G, Tolley S, Turkenburg JP, Christiansen L, Huge-Jensen B, Norskov L, Thim L, Menge U. A serine protease triad forms the catalytic centre of a triacylglycerol lipase. Nature 1990;343:767–770.
7. Aghajari N, Feller G, Gerday C, Haser R. Crystal structures of the psychrophilic α-amylase from *Alteromonas haloplanctis* in its native form and complexed with an inhibitor. Protein Sci 1998;7:564–572.
8. Schrag JD, Li YG, Wu S, Cygler M. Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. Nature 1991;351:761–764.
9. Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, Toker L, Silman I. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. Science 1991;253:872–879.
10. Blow D. Enzymology. More of the catalytic triad. Nature 1990;343:694–695.
11. Wallace AC, Laskowski RA, Thornton JM. Derivation of 3D coordinate templates for searching structural databases: application to Ser-His-Asp catalytic triads in the serine proteinases and lipases. Protein Sci 1996;5:1001–1013.
12. Rye CS, Withers SG. Glycosidase mechanisms. Curr Opin Chem Biol 2000;4:573–580.
13. MacGregor EA, Janecek S, Svensson B. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochim Biophys Acta 2001;1546:1–20.
14. Vasella A, Davies GJ, Bohm M. Glycosidase mechanisms. Curr Opin Chem Biol 2002;6:619–629.
15. Janecek S. Alpha-amylase family: molecular biology and evolution. Prog Biophys Mol Biol 1997;67:67–97.
16. Levitzki A, Steer ML. The allosteric activation of mammalian α-amylase by chloride. Eur J Biochem 1974;41:171–180.
17. Lifshitz R, Levitzki A. Identity and properties of the chloride effector binding site in hog pancreatic α-amylase. Biochemistry 1976;15:1987–1993.
18. Qian M, Haser R, Payan F. Structure and molecular model refinement of pig pancreatic α-amylase at 2.1 Å resolution. J Mol Biol 1993;231:785–799.
19. Brayer GD, Luo Y, Withers SG. The structure of human pancreatic α-amylase at 1.8 Å resolution and comparisons with related enzymes. Protein Sci 1995;4:1730–1742.
20. Ramasubbu N, Paloth V, Luo Y, Brayer GD, Levine MJ. Structure of human salivary α-amylase at 1.6 Å resolution: implications for its role in the oral cavity. Acta Crystallogr D 1996;52:435–446.
21. Strobl S, Maskos K, Betz M, Wiegand G, Huber R, Gomis-Ruth FX, Glockshuber R. Crystal structure of yellow meal worm α-amylase at 1.64 Å resolution. J Mol Biol 1998;278:617–628.
22. Aghajari N, Feller G, Gerday C, Haser R. Structures of the psychrophilic *Alteromonas haloplanctis* α-amylase give insights into cold adaptation at a molecular level. Structure 1998;6:1503–1516.
23. Feller G, le Bussy O, Houssier C, Gerday C. Structural and functional aspects of chloride binding to *Alteromonas haloplanctis* α-amylase. J Biol Chem 1996;271:23836–23841.
24. Aghajari N, Feller G, Gerday C, Haser R. Structural basis of α-amylase activation by chloride. Protein Sci 2002;11:1435–1441.
25. Numao S, Maurus R, Sidhu G, Wang Y, Overall CM, Brayer GD, Withers SG. Probing the role of the chloride ion in the mechanism of human pancreatic α-amylase. Biochemistry 2002;41:215–225.
26. Qian M, Ajandouz el H, Payan F, Nahoum V. Molecular basis of the effects of chloride ion on the acid-base catalyst in the mechanism of pancreatic α-amylase. Biochemistry 2005;44:3194–3201.
27. D'Amico S, Gerday C, Feller G. Structural similarities and evolutionary relationships in chloride-dependent α-amylases. Gene 2000;253:95–105.
28. Da Lage JL, Feller G, Janecek S. Horizontal gene transfer from Eukarya to bacteria and domain shuffling: the α-amylase model. Cell Mol Life Sci 2004;61:97–109.
29. D'Amico S, Gerday C, Feller G. Structural determinants of cold adaptation and stability in a large protein. J Biol Chem 2001;276:25791–25796.

30. D'Amico S, Sohier JS, Feller G. Kinetics and energetics of ligand binding determined by microcalorimetry: insights into active site mobility in a psychrophilic α-amylase. J Mol Biol 2006;358:1296–1304.
31. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
32. Barker JA, Thornton JM. An algorithm for constraint-based structural template matching: application to 3D templates with statistical analysis. Bioinformatics 2003;19:1644–1649.
33. Wallace AC, Borkakoti N, Thornton JM. TESS: a geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases. Application to enzyme active sites. Protein Sci 1997;6:2308–2323.
34. Porter CT, Bartlett GJ, Thornton JM. The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. Nucleic Acids Res 2004;32:D129–D133.
35. D'Amico S, Gerday C, Feller G. Temperature adaptation of proteins: engineering mesophilic-like activity and stability in a cold-adapted α-amylase. J Mol Biol 2003;332:981–988.
36. Feller G, d'Amico D, Gerday C. Thermodynamic stability of a cold-active α-amylase from the Antarctic bacterium *Alteromonas haloplanctis*. Biochemistry 1999;38:4613–4619.
37. Baskakov I, Wang A, Bolen DW. Trimethylamine-*N*-oxide counteracts urea effects on rabbit muscle lactate dehydrogenase function: a test of the counteraction hypothesis. Biophys J 1998;74:2666–2673.
38. Hibbert F, Emsley J. Hydrogen bonding and chemical reactivity. Adv Phys Organ Chem 1990;26:255–379.
39. Neal S, Nip AM, Zhang H, Wishart DS. Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts. J Biomol NMR 2003;26:215–240.
40. Frey PA, Whitt SA, Tobin JB. A low-barrier hydrogen bond in the catalytic triad of serine proteases. Science 1994;264:1927–1930.
41. Robillard G, Shulman RG. High resolution nuclear magnetic resonance study of the histidine—aspartate hydrogen bond in chymotrypsin and chymotrypsinogen. J Mol Biol 1972;71:507–511.
42. Tyukhtenko SI, Litvinchuk AV, Chang CF, Leu RJ, Shaw JF, Huang TH. NMR studies of the hydrogen bonds involving the catalytic triad of *Escherichia coli* thioesterase/protease I. FEBS Lett 2002;528:203–206.
43. Seavey BR, Farr EA, Westler WM, Markley JL. A relational database for sequence-specific protein NMR data. J Biomol NMR 1991;1:217–236.
44. Mildvan AS, Harris TK, Abeygunawardana C. Nuclear magnetic resonance methods for the detection and study of low-barrier hydrogen bonds on enzymes. Methods Enzymol 1999;308:219–245.
45. Yufeng W, McDermott AE. Effects of hydrogen bonding on chemicals shifts. In: Facelli JC, De Dios AC, editors. Modeling NMR chemicals shifts: gaining insights into structure and environment. ACS Books; 1999. pp177–193.
46. Rosen MK, Yamazaki T, Gish GD, Kay CM, Pawson T, Kay LE. Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. Nature 1995;374:477–479.
47. Charron C, Talfournier F, Isupov MN, Littlechild JA, Branlant G, Vitoux B, Aubry A. The crystal structure of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon *Methanothermus fervidus* in the presence of NADP(+) at 2.1 Å resolution. J Mol Biol 2000;297:481–500.

 DOI 10.1002/prot