Cloning of Amadoriase I Isoenzyme from *Aspergillus sp.*: Evidence of FAD Covalently Linked to Cys342[†,‡]

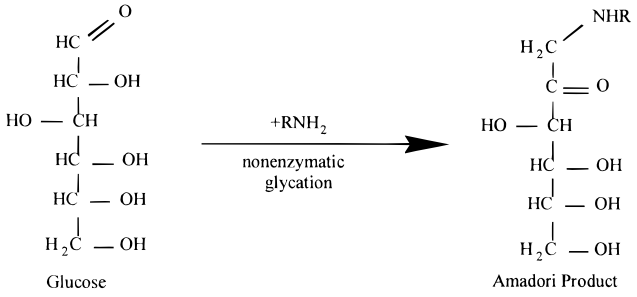
Xinle Wu,[§] Motoko Takahashi,[§,||] Shu G. Chen, and Vincent M. Monnier*

Institute of Pathology, Case Western Reserve University, Cleveland, Ohio 44106

Received August 30, 1999; Revised Manuscript Received November 22, 1999

ABSTRACT: Amadoriases are a novel class of FAD enzymes which catalyze the oxidative deglycation of glycated amino acids to yield corresponding amino acids, glucosone, and H_2O_2. We previously reported the purification and characterization of two amadoriase isoenzymes from *Aspergillus fumigatus* and the molecular cloning of amadoriase II. To identify the primary structure of amadoriase I, we prepared a cDNA library from *Aspergillus fumigatus* and isolated a clone using a probe amplified by polymerase chain reaction with primers designed according to the partial amino acid sequences from peptide mapping. The primary structure of the enzyme deduced from the nucleotide sequence comprises 445 amino acid residues. The enzyme contains 1 mol of FAD as a cofactor, which is covalently linked to Cys342, as determined by mutagenesis analysis, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, and electrospray ionization−collisional-activated dissociation tandem mass spectrometry. Sequence alignment studies show that amadoriase I has 22% homology with monomeric sarcosine oxidase in which FAD is also linked to a homologous Cys residue. Amadoriases are of potential importance as tools for uncoupling hyperglycemia and glycation reactions that are thought to play a role in diabetic complications.

Nonenzymatic glycation of proteins has been implicated in the development of diabetic complications and the aging process (*1*). In this reaction, glucose reacts nonenzymatically with amino groups in proteins to form ketoamine-linked 1-deoxyfructosyl adducts:



The Amadori product, which forms in the initial stages of the Maillard reaction in vivo, is the precursor of protein cross-links, fluorescent, UV-active compounds, and glycoxidation products which accumulate in long-lived proteins with time (*2*−*4*). It is also a potential source of reactive oxygen species (*5*, *6*) and can be taken up by specific receptors on macrophages, mesangial, and aortic cells (*7*−*9*).

The work was originally initiated with the concept in mind that uncoupling of glycation from hyperglycemia in transgenic animal models of diabetes would be helpful to specifically study the role of protein glycation in diabetic complications since the enzyme can regenerate free lysine amino groups. Previously, we isolated and partially characterized two amadoriase[1] isoenzymes from an *Aspergillus fumigatus* from soil which deglycate Amadori products to yield corresponding amino acids, glucosone, and H_2O_2 (*10*). This and other deglycating enzymes, also called fructosyl amino acid oxidases, have FAD as a cofactor (*11*−*14*). In this paper, we describe the cloning, the expression of cDNA of amadoriase I from a cDNA library of *Aspergillus fumigatus*, and the identification of cysteine 342 as the covalent site of attachment of FAD.

EXPERIMENTAL PROCEDURES

Materials. Fructosyl propylamine was prepared as described previously (*10*). Acromobactor protease was purchased from Wako Chemical (Osaka, Japan). Endoproteinase Glu-C was obtained from Sigma (St. Louis, MO). Restriction enzymes were purchased from New England Biolabs (Beverly, MA) and Gibco (Grand Island, NY). T4 DNA ligase

[†] This work was supported in part by Grant EY 07099 from the National Eye Institute. M.T. was supported by a Mentorship grant (to V.M.M.) from the American Diabetes Association and was a Fellow of the Juvenile Diabetes Foundation International.

[‡] The nucleotide sequence reported in this paper has been submitted to the GenBank with accession number AF035700.

* To whom correspondence should be addressed. Fax: (216)368-0495. Phone: (216)368-6613. Email: vmm3@po.cwru.edu.

[§] The first two authors contributed equally to this study and are therefore co-first authors.

[||] Present address: Department of Biochemistry, Osaka University Medical School, Suita 565, Japan.

[1] Abbreviations: amadoriase, fructosyl amino acid oxidase (EC 1.3.5); CAD, collisional activated dissociation; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; PCR, polymerase chain reaction; HPLC, high-performance liquid chromatography; dNTP, deoxynucleotide triphosphate; dCTP, deoxycytidine triphosphate; IPTG, isopropyl-D-thiogalactoside; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; OPD, *o*-phenylenediamine.

10.1021/bi992031g CCC: $19.00
Published on Web 01/19/2000

and Taq DNA polymerase were obtained from Boehringer Mannheim (Indianapolis, IN) and Gibco, respectively. QIAquick Gel Extraction Kit and Ni-NTA Superflow resin were purchased from Qiagen (Valencia, CA). All other materials were analytical grade.

Peptide Mapping and Amino Acid Sequencing. Amadoriase I was purified as described previously (*10*). About 50 μg of the protein was digested with Acromobactor protease and applied to a reverse-phase C18 HPLC column (ODS, 4.6 × 150 mm, Vydac, The Separations Group, Hesperia, CA), and a linear gradient from 0 to 60% acetonitrile in 0.1% trifluoroacetic acid was applied at a flow rate of 1.0 mL/min. Peptides were detected by absorbance measurements at 214 nm. Amino acid sequencing of peptides was performed at the Molecular Biology Core Laboratory (Department of Biochemistry, Case Western Reserve University) with an Applied Biosystems 477A Protein Sequencer (Foster City, CA).

Polymerase Chain Reaction (PCR). Degenerated sense and antisense oligonucleotide primers were designed according to the obtained amino acid sequences. PCR was carried out in 50 μL reaction mixtures containing a pair of primers, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, and 200 mM dNTP. Thirty cycles (94 °C for 1 min, 55 °C for 2 min, 72 °C for 3 min) of PCR were performed using 2.5 units of *Thermus aquaticus (Taq)* polymerase.

Isolation and Sequencing of Amadoriase I cDNA Clones. The amplification products were purified using acrylamide gel electrophoresis and [^{32}P]dCTP to screen *Aspergillus fumigatus* cDNA library constructed in λZAP vector (*15*). pBluescript I SK(−) plasmid containing the cDNA insert from positive bacteriophage isolated was recovered by in vivo excision using helper phage. The sequence of both strands of selected cDNA clones was determined by the dideoxy chain termination method.

Site-Directed Mutagenesis and Cloning of Amadoriase I cDNA into Expression Vector. Oligonucleotide-directed mutagenesis was used to construct cDNA sequences encoding cysteine to serine or alanine substitution at position 342. Site-directed mutagenesis was carried out using the Altered Site II in vitro Mutagenesis Systems from Promega. Amadoriase I sequences were amplified by PCR with primers which will introduce *Bam*HI and *Nde*I sites at the ends. A resulting *Bam*HI−*Nde*I fragment was digested with both restriction enzymes and ligated into expression vector pET 15b (Novagen, Madison, WI).

Expression and Purification of Recombinant Amadoriase I. Plasmids carrying wild-type and mutant amadoriase I cDNAs were transformed into *E. coli* strain BL21(DE3)-pLysS. The cells were harvested 4−5 h after induction with 0.25 mM isopropyl-β-D-thiogalactoside (IPTG) and frozen at −70 °C. Cells were resuspended in lysis buffer (50 mM NaH_2PO_4, 300 NaCl, 10 mM imidazole, pH 8.0) and broken by four, 30 s sonications. Membrane fragments were removed by centrifugation, and whole cell extracts were applied to an FPLC column (Pharmacia) containing Ni-NTA superflow resin equilibrated in lysis buffer. The column was washed with wash buffer (50 mM NaH_2PO_4, 300 NaCl, 20 mM imidazole, pH 8.0) at a flow rate of 2.0 mL/min until the absorbance at 260 nm stabilized. Proteins were then eluted with elute buffer (50 mM NaH_2PO_4, 300 NaCl, 200 mM imidazole, pH 8.0) at a flow rate of 2.0 mL/min. Protein concentration was determined by the method of Bradford (*16*) with a Bio-Rad Protein Assay kit (Bio-Rad Laboratories) using bovine serum albumin as a standard. Protein analysis by SDS−PAGE was carried out according to Laemmli (*17*). Approximately 30−40 mg of purified amadoriase I was obtained from 2 L of induced cells. Proteins were dialyzed against 50 mM ammonium bicarbonate (pH 7.8) overnight before the proteolytic digestion was performed.

Assay for Enzymatic Activity. The enzyme activity was monitored by the release of glucosone measured by a colorimetric reaction with *o*-phenylenediamine (OPD) as described previously (*10*). Briefly, the enzyme solution was added to the reaction mixture containing 20 mM sodium phosphate, pH 7.4, 10 mM OPD, and 10 mM fructosyl propylamine in a final volume of 1 mL. After incubation at 37 °C for 2 h, the absorbance at 320 nm was measured. One unit of enzyme activity was defined as the amount of enzyme that produces 1μmol of glucosone per minute. Synthetic glucosone was used as a standard.

Proteolytic Digestion of Wild-Type or Mutant Amadoriase I and Isolation of the FAD Peptide. Samples of recombinant amadoriase I (0.6 mg/mL) were digested with endoproteinase Glu-C using a ratio (w/w) of 1:50 in 50 mM ammonium bicarbonate (pH 7.8) at 37 °C for about 24 h. The proteolytic digest was analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) to obtain sequence-specific information on peptides and putative modifications. To isolate the flavin peptide from the wild-type preparation, the Glu-C digest was applied to a C_{18} reverse-phase HPLC column (4.6 × 250 mm, Vydac). Elution profile (flow rate = 1.0 mL/ml): 40 min linear gradient from 0 (0.036% trifluoroacetic acid) to 100% acetonitrile (0.045% trifluoroacetic acid). The column eluate was monitored by its absorbance at 214 nm and fluorescence (λ_{ex} = 452 nm, λ_{em} = 525 nm). The fraction containing the FAD peptide was collected and freeze-dried.

Amadoriase was also digested with the modified trypsin (sequencing grade, Boehringer) at the enzyme-to-substrate ratio of 1:100 for 18 h at 37 °C. The tryptic peptides were purified on a reverse-phase C18 column (3 Å; 1 × 150 mm) with a gradient of 5−70% buffer B (98% acetonitrile, 0.01% TFA) at a flow rate of 50 μL/mL using a HPLC system (Michrom Bioresources, Auburn, CA). The column was monitored on-line by UV absorbance at 450 and 214 nm. Flavin-containing peptides were initially located by their high absorbance at 450 nm. They were then characterized by either MALDI or electrospray ionization mass spectrometry.

Mass Spectral Analysis. The polypeptides were initially analyzed by MALDI-TOF-MS. The spectrometer was equipped with a pulsed nitrogen laser source (λ = 337 nm) and used an acceleration voltage of 28 kV operated in the linear mode (Voyager Biospectrometry Workstation, PerSeptive Biosystems, Framingham, MA). Samples were mixed with an equal volume of saturated α-cyano-4-hydroxycinnamic acid solution in 60% acetonitrile/0.3% trifluoroacetic acid. One microliter of mixture was applied onto the laser target probe and was air-dried before being introduced into the mass spectrometer. Between 100 and 200 spectra were obtained for each sample. They were averaged and analyzed using the GRAMS software on the instrument. The instrument was calibrated using a mixture of standard peptides. A mass accuracy ranging from 0.01 to 0.1% was routinely

```
   1 gtccaatccgcgttgccagcgtaaaacactgcctcaaactggttca ATG GCG CCT TCA ATT TTG AGC ACT   70
   1                                                 M   A   P   S   I   L   S   T    8

  71 GAA TCT TCC ATT ATC GTT ATC GGA GCA GGC ACA TGG GGC TGC TCA ACT GCT CTG CAC CTC  130
   9 E   S   S   I   I   V   I   G   A   G   T   W   G   C   S   T   A   L   H   L   28

 131 GCT CGT CGA GGC TAC AAA GAT GTC ACT GTT CTC GAC CCT CAT CCA GTT CCT TCG CCC ATT  190
  29 A   R   R   G   Y   K   D   V   T   V   L   D   P   H   P   V   P   S   P   I   48

 191 GCA GCA GGC AAT GAC ATC AAC AAG ATT ATG GAG CAC AGC GAG CTG AAA GAT GGC TCA TCC  250
  49 A   A   G   N   D   I   N   K   I   M   E   H   S   E   L   K   D   G   S   S   68

 251 GAC CCT CGA AGC GCA GCC TTC TCG ACA TTT ACG CGA GCT GCT CTT AAG GCG TGG AAA ACT  310
  69 D   P   R   S   A   A   F   S   T   F   T   R   A   A   L   K   A   W   K   T   88

 311 GAC CCG GTT TTC CAG CCT TAC TTT CAC GAA ACT GGC TTT ATC ATA TCG GGG CAC ACA CCT  370
  89 D   P   V   F   Q   P   Y   F   H   E   T   G   F   I   I   S   G   H   T   P  108

 371 GCT CTG ATT GAC CAC ATA CGA AAA GAC GAG GTA GAA CCG TCA GAA ACA AAC TTC GTC AAG  430
 109 A   L   I   D   H   I   R   K   D   E   V   E   P   S   E   T   N   F   V   K  128

 431 CTG GAG ACA GCC GAG GAC TTC CGC CGG ACC ATG CCG CCA GGT GTA CTG ACA GGC GAC TTC  490
 129 L   E   T   A   E   D   F   R   R   T   M   P   P   G   V   L   T   G   D   F  148

 491 CCT GGC TGG AAA GGC TGG TTG CAC AAG TCT GGT GCT GGG TGG ATT CAT GCC AAA AAG GCT  550
 149 P   G   W   K   G   W   L   H   K   S   G   A   G   W   I   H   A   K   K   A  168

 551 ATG ATC TCT GCT TTC AAT GAA GCT AAG CGC TTG GGA GTC AGA TTT GTC ACT GGC TCT CCG  610
 169 M   I   S   A   F   N   E   A   K   R   L   G   V   R   F   V   T   G   S   P  188

 611 GAA GGG AAT GTT GTA TCG TTG GTA TAC GAG GAC GGA GAC GTC GTT GGA GCC AGA ACT GCC  670
 189 E   G   N   V   V   S   L   V   Y   E   D   G   D   V   V   G   A   R   T   A  208

 671 GAT GGT CGC GTG CAC AAA GCC CAT CGC ACT ATT CTT TCG GCA GGT GCT GGC AGT GAC AGT  730
 209 D   G   R   V   H   K   A   H   R   T   I   L   S   A   G   A   G   S   D   S  228

 731 CTC CTA GAC TTC AAG AAG CAG CTT CGG CCT ACC GCG TGG ACT CTC TGT CAT ATT CAG ATG  790
 229 L   L   D   F   K   K   Q   L   R   P   T   A   W   T   L   C   H   I   Q   M  248

 791 GGC CCT GAA GAG GTC AAG CAA TAT CGG AAC CTT CCT GTG TTG TTC AAC ATC GCC AAA GGG  850
 249 G   P   E   E   V   K   Q   Y   R   N   L   P   V   L   F   N   I   A   K   G  268

 851 TTC TTC ATG GAG CCT GAT GAG GAT AAA CAC GAG CTC AAG ATT TGT GAC GAG CAT CCA GGG  910
 269 F   F   M   E   P   D   E   D   K   H   E   L   K   I   C   D   E   H   P   G  288

 911 TAC TGC AAC TTT CTC CCT GAC CCA AAC AGA CCG GGC CAG GAG AAG AGT GTC CCC TTC GCA  970
 289 Y   C   N   F   L   P   D   P   N   R   P   G   Q   E   K   S   V   P   F   A  308

 971 AAG CAT CAG ATC CCG CTC GAG GCC GAA GCC CGC GCA CGA GAC TTT CTC CAT GAT ACA ATG 1030
 309 K   H   Q   I   P   L   E   A   E   A   R   A   R   D   F   L   H   D   T   M  328

1031 CCG CAT CTG GCT GAC CGG CCA CTG TCT TTC GCG CGT ATT TGC TGG GAT GCT GAT ACC CCA 1090
 329 P   H   L   A   D   R   P   L   S   F   A   R   I   C   W   D   A   D   T   P  348

1091 GAC CGT GCT TTC TTG ATC GAT AGA CAT CCT GAA CAC CCC TCA CTG CTA GTC GCT GTT GGA 1150
 349 D   R   A   F   L   I   D   R   H   P   E   H   P   S   L   L   V   A   V   G  368

1151 GGT TCC GGC AAT GGC GCC ATG CAA ATG CCT ACA ATT GGC GGT TTT ATC GCA GAT GCT CTA 1210
 369 G   S   G   N   G   A   M   Q   M   P   T   I   G   G   F   I   A   D   A   L  388

1211 GAG AGT AAA CTA CAG AAG GAG GTG AAG GAC ATC GTT CGA TGG AGG CCA GAG ACG GCT GTC 1270
 389 E   S   K   L   Q   K   E   V   K   D   I   V   R   W   R   P   E   T   A   V  408

1271 GAT CGA GAT TGG AGA GCG ACT CAG AAT CGC TTT GGC GGG CCT GAC AGG ATC ATG GAT TTT 1330
 409 D   R   D   W   R   A   T   Q   N   R   F   G   G   P   D   R   I   M   D   F  428

1331 CAG CAG GTC GGA GAG GAT CAG TGG ACC AAG ATT GGA GAG AGC AGA GGT CCG TAA tagacctt 1392
 429 Q   Q   V   G   E   D   Q   W   T   K   I   G   E   S   R   G   P   *           446

1393 gcggcgtgttcgccacattgaaggctagcatatcttgaggagaagtcgtctagacaacatgttttggagctagccacagt 1472

1473 actttgagagttatgaagagactcactgttctgctcaaaaaaaaaaaaaaaaaa                          1526
```

FIGURE 1: Nucleotide and amino acid sequences of *Aspergillus fumigatus* amadoriase I. The stop codon is indicated by an asterisk. Residues corresponding to peptide sequences obtained from isolated protein by Edman degradation are underlined (dashed lines indicate unidentified residues).

obtained. Electrospray mass spectral analysis was performed at the Cleveland Mass Spectrometry Facility on a Micromass Quattro II triple quadrupole mass spectrometer (Manchester, U.K.) interfaced with nitrogen as the nebulizer gas and an ionspray voltage of ~3500 V. The sample was infused into the mass spectrometer at 3 μL/min using a Harvard Apparatus syringe pump after dissolving in 20% acetonitrile/0.3% formic acid. Collisional-activated dissociation (CAD)

FIGURE 2: Sequence alignment of homologous proteins. The covalent flavin attachment site in amadoriase I is marked by an asterisk (this study).

was performed using argon as the collision gas at a pressure of 2.7e-3 mBar and a collision energy of 17.0 eV. Under these CAD conditions, product ion spectra of the selected precursor ions ranging from 50 to 2000 were obtained.

RESULTS

Molecular Cloning of Amadoriase I. Amadoriase I protein was purified from *Aspergillus fumigatus* induced with fructosyl propylamine. An HPLC peptide map of the enzyme digested with Acromobactor protease was generated, and the peptides obtained were partially sequenced. The N-terminal sequence and the sequence of the internal peptide eluting at 47.8 min were used to design PCR primers. Acrylamide gel electrophoresis of the PCR product showed a 1.2 kb DNA fragment (data not shown). The PCR product was then cut out of the gel, eluted, labeled with [^{32}P]dCTP, and used as a probe for screening a UniZAP cDNA library of *Aspergillus fumigatus* prepared as described previously (*15*). Upon screening 5×10^4 plaques, 1 clone was isolated and sequenced. The sequence showed an open reading frame of 1381 bp, which encodes a 445 amino acid polypeptide with a predicted molecular mass of 50 kDa. The N-terminal and internal peptide amino acid sequences are included within the predicted sequence, as underlined in Figure 1. Sequence alignments with existing amadoriases and other related proteins are shown in Figure 2. Amadoriase I shows 51% sequence identity with amadoriase II and 37% and 52% with fructosyl amino acid oxidase from *P. janthinellum* and *A. terreus*, respectively.

Mutagenesis. The amadoriase I cDNA sequence on pBluescript plasmid was cloned to pAlter-1 vector at restriction sites *Kpn*I and *Pst*I. Altered Site II in vitro Mutagenesis System was used to generate the mutations C342A and C342S. This site was tentatively identified as being part of the active site based on the fact that it is highly conserved among existing fructosyl amino acid oxidases, and on sequence homology with sarcosine oxidase, a related enzyme having Cys315 in the active site (*18*). Mutants were amplified by PCR with primers designed to introduce the restriction sites *Bam*HI and *Nde*I at the ends. PCR products were digested with these two enzymes and ligated into the expression vector pET15b which will generate a fusion protein with a histidine tag at the N-terminus of the enzyme. Amadoriase I sequences were completely sequenced after cloning into expression plasmids to ensure that no unintended mutations had been introduced. The final constructs were transformed into *E. coli* strain BL21(DE3)pLysS. The expression level of the mutant enzymes was similar to the wild-type recombinant enzyme after inducing with IPTG base on the band intensity on SDS−PAGE. Proteins were purified by Ni-NTA affinity column.

Activity. Apparent kinetic constants of wild-type and mutant amadoriase I for fructosyl propylamine are shown in Table 1. The mutants lost most of the activity compared to the wild type.

Identification of the Covalent Flavin Attachment Site. Samples of wild-type and mutant recombinant amadoriase I

Table 1: Kinetic Constants (Wild-Type vs Mutants)[a]

	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} min^{-1})
WT	0.049	118	2.4×10^6
C342A	0.36	16	4.4×10^4
C342S	0.28	5	1.7×10^4

[a] Activity was measured by the glucosone formation assay as described under Experimental Procedures. Apparent kinetic constants were obtained by least-squares linear regression analysis of the data obtained under nonsaturating conditions.

were subjected to proteolytic digestion with endoproteinase Glu-C. Peptides derived from digestion of wild-type, C342A, and C342S mutants were analyzed by MALDI-MS (Figure 3). Assignment of peptides to their sequences was made according to the observed and expected mass values of protonated ions, and the known cleavage pattern of Glu-C. MALDI spectra of the whole digests from the wild-type, C342A, and C342S enzymes revealed many peptides with the same [M+H$^+$] values corresponding to the common regions of the enzymes (data not shown), as well as some peptides with different mass values. Only the mass region showing such differences was presented for brevity (Figure 3). For the wild type, peptide 318−389 with expected [M+H$^+$] values at 7840.9 was identified only as a minor signal. Instead, two prominent signals at 8624.0 and 8295.9 were found (Figure 3, top panel), which corresponded in mass values to those of peptide 318−389 linked to FAD and FMN, respectively. For C342A (Figure 3, middle panel) and C342S (Figure 3, bottom panel), only unmodified peptide 318−389 was recognized at [M+H$^+$] values of 7807.9 and 7823.4, respectively. These data are consistent with Cys342 being the attachment site of flavin (Table 2). When the Glu-C digest from the wild-type enzyme was analyzed by HPLC with fluorescence detection (λ_{ex} = 452 nm, λ_{em} = 525 nm), a major peptide peak at 22.1 min with spectral properties of flavin was found (data not shown). The collected HPLC peak was subsequently identified as peptide 318−389 linked to FAD and FMN by MALDI-MS. However, no flavin-linked peaks were found to elute from HPLC for the digest of both mutants. Thus, the replacement of Cys342 with either Ala or Ser abolished modification by flavin. It is likely that FAD rather than FMN is initially linked to Cys342. However, since FAD is light sensitive and unstable during prolonged storage (unpublished observation), degradation of the FAD could occur during the experimental procedures, resulting in the appearance of a low amount (~30%) of the FMN-linked peptide. Alternatively, the high expression level in *E. coli* may have led to incomplete incorporation of FAD into the enzyme (*19*).

The results of the mutation and MALDI experiments suggested but did not prove that FAD was linked to Cys342, because the mutation can affect the structure and inhibit flavination at other sites. For example, our enzyme contains a conserved His at residue 357 which could have been the site of a covalent linkage. To further exclude this possibility, we have digested the enzyme with trypsin and isolated flavin-containing peptides by collecting the HPLC peak with the highest A_{450nm}/A_{214nm} ratio and analyzed it by MALDI-MS (Figure 4). A FAD-modified peptide at [M+H$^+$] 1977.6 was found, corresponding to that predicted for peptide 341−350 linked to FAD. The FMN-modified peptide, likely to be a product of partial hydrolysis that releases AMP from the FAD moiety, and the unmodified peptide were also detected at much lower intensities.

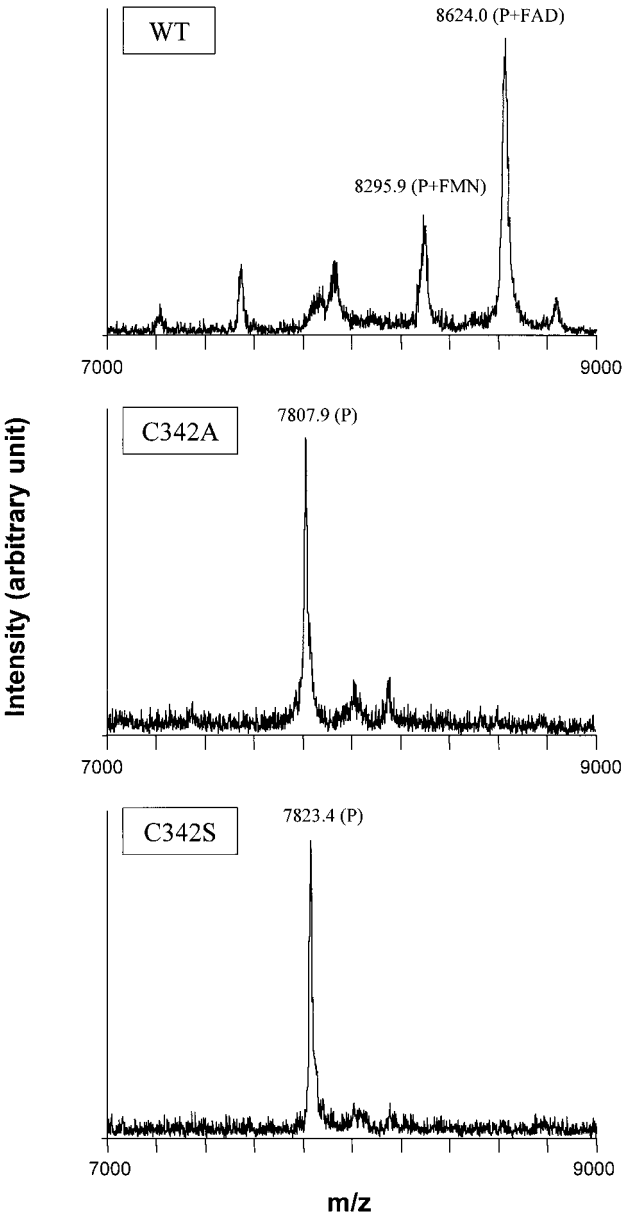


FIGURE 3: Identification of Cys342 as the putative site of flavin modification. Peptides (P) derived from endoproteinase Glu-C digestion of amadoriase I wild-type (WT) and C342A and C342S mutants were analyzed by MALDI-MS. All data were expressed as average mass values. For brevity, only the region (*m*/*z* 7000−9000) showing differences in peptide mass values between the three enzymes is shown. For the wild-type (upper panel), [M+H$^+$] values (observed vs expected in parentheses) at 8624.0 (8624.9) and 8295.9 (8296.9) corresponded to peptide 318−389 (expected [M+H$^+$] for the unmodified peptide: 7840.9) linked to FAD and FMN, respectively. For C342A (middle panel) and C342S (lower panel) mutants, only unmodified peptide 318−389 was recognized at [M+H$^+$] values of 7807.9 (7808.8) and 7823.4 (7824.8), respectively.

Table 2: MALDI-MS Analysis of Peptide from Wild-Type and C342A and C342S Amadoriase I

source of peptide	[M+H$^+$] values observed	predicted	sequence	modification
wild-type	8624.0	8624.9	318−389 (wild-type)	FAD
wild-type	8295.9	8296.9	318−389 (wild-type)	FMN
C342A	7807.9	7808.8	318−389 (C342A)	no
C342S	7823.4	7824.8	318−389 (C342S)	no

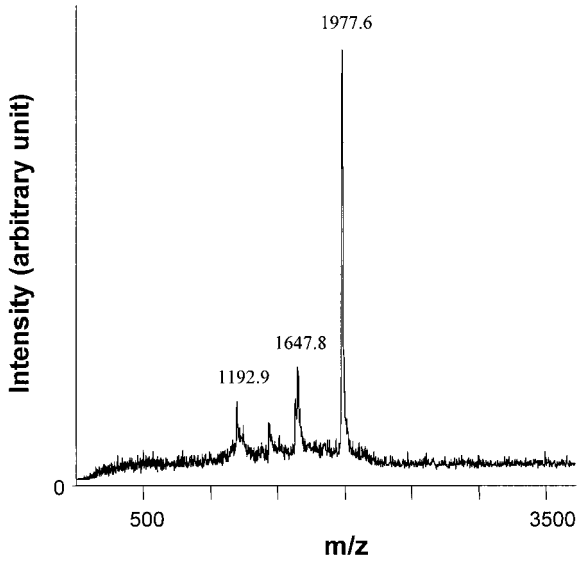


FIGURE 4: MALDI-MS spectrum of a tryptic peptide modified by FAD. A HPLC fraction containing the FAD-modified peptide detected by on-line UV absorbance at 450 nm was analyzed by MALDI-MS. A major signal with a $[M+H^+]$ value of 1977.6 corresponded to that of tryptic peptide 341−350 linked to FAD (expected $[M+H^+]$: 1976.3). The signals with lower intensities at $[M+H^+]$ values of 1647.8 and 1192.9 corresponded to those of the FMN-modified (expected $[M+H^+]$: 1648.3) and unmodified (expected $[M+H^+]$: 1192.3) peptides, respectively.

Since in the tryptic peptide 341−350 Cys342 is the only available site of modification (other residues known to be modified by flavin are histidine and tyrosine), it is important to confirm the initial result based on the MALDI-MS analysis. More definitive evidence for the assignment of this peptide and the site of FAD attachment was provided by amino acid sequence analysis and by collisional-activated dissociation (CAD) tandem MS. The HPLC peak containing FAD-modified tryptic peptide was subjected to N-terminal sequencing by automated Edman degradation. An amino acid sequence of I-X-W-D-A-D-T-P-D-R, corresponding to residues 341−350 of amadoriase I, was obtained (Table 3). No amino acid derivative was detected in the second cycle of the sequence which coincides with Cys342. To confirm that FAD is indeed covalently linked to this residue, we further characterized the peptide using electrospray ionization−CAD tandem MS on a triple quadrupole instrument. A triply charged precursor ion at *m*/*z* 659 in the first quadrupole was selected for collision with argon in the second quadruple; fragmented product ions were analyzed in the third quadrupole, producing a characteristic CAD spectrum of charged product ions (Figure 5). AMP and the doubly charged peptide without AMP were detected, consistent with the presence of a covalently bound FAD moiety which undergoes cleavage at the phosphodiester bond. The observed N-terminal fragments (a_1,b_2-AMP, b_4-AMP, b_5-AMP, b_6-AMP, b_7-AMP, and

Table 3: Amino Acid Sequence Analysis of the FAD-Containing Peptide Purified from a Tryptic Digest of Amadoriase I

cycle	residue	yield (pmol)
1	I	389.96
2	X[a]	—[b]
3	W	263.45
4	D	241.66
5	A	278.34
6	D	228.05
7	T	155.71
8	P	162.70
9	D	125.55
10	R	157.61

[a] No amino acid derivative was obtained at this cycle. [b]Similar to that of background.

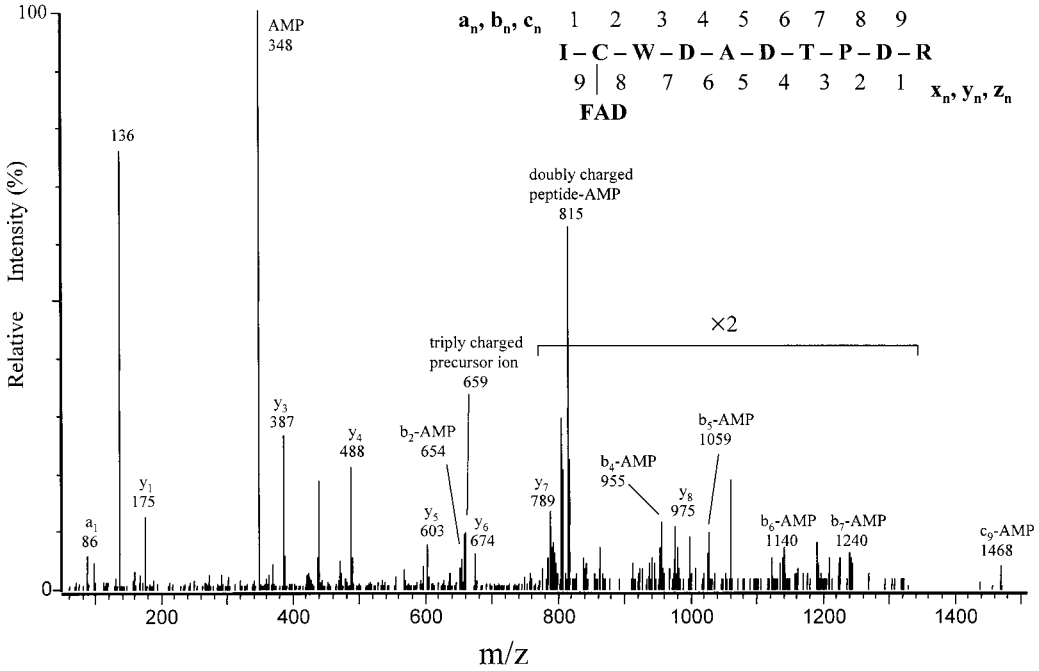


FIGURE 5: CAD mass spectrum of the FAD peptide isolated from a tryptic digest of amadoriase I. The triply charged ion of the FAD peptide (659 Da) was selected as the precursor ion. The peptide sequence and numbering for N-terminal (a_n, b_n, c_n) and C-terminal (x_n, y_n, z_n) fragments are listed. The three subtypes of fragments represent the products of cleavage between C(α) and C═O (a_n, x_n), C═O and NH (b_n, y_n), or NH and C(α) (c_n, z_n). An ion with −AMP (minus AMP) indicates that AMP is lost from the FAD moiety due to cleavage at the phosphodiester linkage.

C_9-AMP) suggest that the Cys residue is modified by flavin while other residues are unmodified. In addition, when combined with the C-terminal fragments (y_1, y_3, y_4, y_5, y_6, y_7, y_8), the CAD spectral data provide further support for the proposed sequence of the peptide with no modification at residues other than cysteine. Therefore, our results identify Cys342 as the attachment site for FAD. Furthermore, spectroscopic analysis showed a typical 452 nm absorption spectrum (*10* and data not shown) compatible with 8-α-cysteinyl rather than 6-α-cysteinyl linkage (*21*). Taken together, the present study strongly supports the presence of 8-α-cysteinyl-linked FAD at position 342 in amadoriase I.

DISCUSSION

We have cloned and sequenced amadoriase I from a cDNA library from *Aspergillus sp*. This enzyme has 51% sequence identity with amadoriase II which we cloned in the previous study, and 34% and 50% identity to two other fructosyl amino acid oxidases from *P. janthinellum* and *A. terreus*, respectively. All these enzymes have a conserved cysteine residue located in the C-terminal region which is Cys342 for amadoriase I (Figure 2). A search for protein homologues using BLAST reveals that monomeric sarcosine oxidase from *Bacillus sp*., which also has the conserved cysteine residue (Cys315), has 22% identity with amadoriase I. The *Bacillus sp*. enzyme contains 1 mol of FAD which is bound covalently through an 8-α-*S*-cysteinyl linkage to Cys315 (*20*). We mutated cysteine 342 of amadoriase I to alanine and serine and found that the enzyme had lost most of its activity. We interpret the residual activity of 5% as being due to trapping of noncovalently bound FAD. Using MALDI-TOF and electrospray−CAD mass spectrometry, we demonstrated that flavin is also covalently linked to amadoriase I through Cys342. According to the absorption spectrum of the amadoriase enzyme, which shows two major absorption bands at 362 and 452 nm (*10*), most likely FAD is also bound to cysteine 342 through an 8-α-*S*-cysteinyl linkage (*21*).

Interestingly, following purification, the mutant enzymes still had two absorption maxima at 362 and 452 nm and residual activity (Table 1), suggesting that FAD was still incorporated into the enzymes during the expression. This suggests that the FAD binding cavity, such as the ADP binding $\beta\alpha\beta$-fold near the N-terminus of the enzyme (*22*), is still intact after the mutation of the cysteine and that FAD was noncovalently bound. Approximately 80% and 50% of FAD, on a mole per mole basis, were noncovalently bound in the C342A and C342S mutants, respectively. Thus, the fact that the mutants lost most of the activity compared to the wild type suggests the covalent linkage between amadoriase I and FAD cofactor must play an important role in the catalytic function of the enzyme, such as fixing the orientation of the cofactor.

The enzyme catalyzes the reaction using flavin as a cofactor, and the overall catalytic cycle consists of two stages. In the first stage, the covalently linked flavin (FAD) moiety is likely converted to its reduced form ($FADH_2$) in concert with oxidation of the 1,2-imine intermediate which is the precursor of the free aminolysine and the aldehyde of glucosone. Regeneration of the oxidized form of the enzyme then occurs in the second step by the action of molecular oxygen. In many reactions in which flavin oxidizes a reactant with an acidic carbon, the enol of this carbon acid is involved as an intermediate. So reducing the carbonyl oxygen on the Amadori product to a hydroxy group might generate a potential inhibitor of the amadoriase. Surprisingly, we found the synthetic borohydride-reduced substrate 1-deoxyglucitolyl propylamine does not inhibit the enzyme activity, which suggests that the carbonyl group might participate in forming a pyranose or furanose ring structure which could serve for potential recognition by the enzyme. The mechanism of the oxidation by the flavin cofactor might, by analogy with sarcosine oxidase, involve the interaction between nitrogen and the flavin ring via electron transfer, hydride transfer, or the formation of a covalent flavin−substrate adduct intermediate (*20*). Further studies on the enzyme mechanism using chemical methods and crystallography are in progress.

ACKNOWLEDGMENT

We thank Cheryl Owens at the Molecular Biology Core for assistance with peptides and DNA sequencing, Vladimir Capka at the Cleveland Mass Spectrometry Facility for the electrospray−CAD experiment, Drs. Witold Surewicz, Wieslaw Swietnicki, and Manuel Morillas for useful discussion, and Dr. Paul Carey for critical reading of the manuscript.

REFERENCES

1. Monnier, V. M. (1990) *J. Gerontol. 45*, B105−111.
2. Ahmed, M. U., Thorpe, S. R., and Baynes, J. W. (1986) *J. Biol. Chem. 261*, 4889−4894.
3. Grandhee, S. K., and Monnier, V. M. (1991) *J. Biol. Chem. 266*, 11649−11653.
4. Dyer, D. G., Blackledge, J. A., Thorpe, S. R., and Baynes, J. W. (1991) *J. Biol. Chem. 266*, 11654−11660.
5. Sakurai, T., and Tsuchiya, S. (1988) *FEBS Lett. 236*, 406−410.
6. Mullarkey, C. J., Edelstein, D., and Brownlee, M. (1990) *Biochem. Biophys. Res. Commun. 173*, 932−939.
7. Salazar, R., Brandt, R., and Krantz, S. (1995) *Biochim. Biophys. Acta 1266*, 57−63.
8. Wu, V.-Y., and Cohen, M. P. (1993) *Biochem. Biophys. Res. Commun. 193*, 1131−1136.
9. Ziyadeh, F. N., and Cohen, M. P. (1993) *Mol. Cell. Biochem. 125*, 19−25.
10. Takahashi, M., Pischetsrieder, M., and Monnier, V. M. (1997) *J. Biol. Chem. 272*, 3437−3443.
11. Horiuchi, T., and Kurokawa, T. (1991) *Agric. Biol. Chem. 55*, 333−338.
12. Horiuchi, T., Kurokawa, T., and Saito, N. (1989) *Agric. Biol. Chem. 53*, 103−110.
13. Yoshida, N., Sakai, Y., Isogai, A., Fukuya, H., Yagi, M., Tani, Y., and Kato, N. (1996) *Eur. J. Biochem. 242*, 499−505.
14. Sakai, Y., Yoshida, N., Isogai, A., Tani, Y., and Kato, N. (1995) *Biosci. Biotechnol. Biochem. 59*, 487−491.
15. Takahashi, M., Pischetsrieder, M., and Monnier, V. M. (1997) *J. Biol. Chem. 272*, 12505−12507.
16. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−54.
17. Laemmli, U. K. (1970) *Nature 227*, 680−685.
18. Wagner, M. A., Khanna, P., and Jorns, M. S. (1999) *Biochemistry 38*, 5588−5595.
19. Packman, L. C., Mewies, M., and Scrutton, N. S. (1995) *J. Biol. Chem. 270*, 13186−13191.
20. Trickey, P., Wagner, M. A., Jorns, M. S., and Mathews, F. S. (1999) *Structure Fold. Des. 7*, 331−345.
21. Decker, K., and Brandsch, R. (1997) *Methods Enzymol. 280*, 413−423.
22. Wierenga, R. K., Terpstra, P., and Hol, W. G. (1986) *J. Mol. Biol. 187*, 101−107.

BI992031G