Inquiro, the UAB undergraduate science research journal

Volume 2007 | Number 1 Article 21

2007

Strutural Characterization of Bacillus anthracis NAD+ Synthetase by Limited Proteolysis

Stephanie J. Hirst

CHEMISTRY

Structural Characterization of *Bacillus anthracis* NAD+ Synthetase by Limited Proteolysis

Stephanie J. Hirst

NAD+ synthetase is the enzyme responsible for the conversion of nicotinic acid adenine dinucleotide (NaAD) into NAD+ in the final step of NAD+ biosynthesis. The prokaryotic and eukaryotic forms of NAD+ synthetase (or NADS) are different enough in size, enzymatic activity, and substrate requirements that the enzyme may serve as a target in antibiotic development. *B. anthracis* NADS was subjected to limited proteolysis by trypsin and chymotrypsin in order to obtain information on the structure of the protein in solution, which was suspected to be different from the crystal structure. The proteolytic fragments were detected by SDS-PAGE and their masses determined by LC-MS. Tryptic cleavage occurred at Lys-9, Lys-86, Leu-95, Lys-132, Lys-206, and Lys-239. Chymotryptic cleavage occurred at Leu-11, Tyr-204, Trp-271, and Leu-277. These sites appear to be susceptible to cleavage because they are located in flexible regions and/or they are highly exposed to solvent. Some cleavages, such as those occurring in the C-terminal helix and at Leu-95, support the hypothesis that the crystal structure differs from the protein in solution because these cleavages would not be expected to occur in rigid secondary structures.

INTRODUCTION

Nicotinamide adenine dinucleotide (NAD+) is a common coenzyme required for several biological functions, including energy metabolism, oxidation-reduction reactions, DNA-repair, and calcium-dependent signal transduction. The ubiquitous coenzyme can be synthesized via two different pathways: 1) a *de novo* pathway, in which NAD+ is synthesized from small molecules, and 2) a pyridine nucleotide salvage pathway. In any case, the last step of NAD+ biosynthesis is catalyzed by NAD+ synthetase (NADS) (1, 2). NADS (E.C. 6.3.5.1) is a common enzyme found in both prokaryotes and eukaryotes, and its catalytic activity is tightly regulated in all organisms (2). It has been shown that it is involved in the synthesis of a variety of biomolecules such as amino acids, purine and pyrimidine dinucleotides, amino sugars, and other coenzymes (1).

NADS is an amidotransferase and has recently been categorized as an "N type" ATP pyrophosphatase, which is a relatively new category of enzyme. The enzymes in this category are thought to share a common mechanism by which substrates are adenylated to promote amidation of the substrate. They often exhibit a "P loop" in the region of the pyrophosphate binding site that is rich in glycine. NADS synthesizes NAD+ from nicotinamide adenine dinucleotide in two steps according to the reactions displayed in Figure 1. In this set of reactions, deamido-NAD^+ is adenylated, and then ammonia is added to the intermediate, resulting in NAD+. This process requires at least two Mg^{2+} ions, which coordinate with active site residues and with ATP, thereby stabilizing the active site (1, 3, 4, 5).

Because the prokaryotic and eukaryotic forms of the enzyme are different in terms of size, enzymatic activity, and substrate requirements, it has recently become a primary target for the development of new antibiotics (3). This is an ever-growing field of research today because there is an unrelenting increase in the number of antibiotic-resistant bacteria, and multidrug resistance is an even more alarming threat (6). Of particular interest in this experiment is the NAD+ synthetase from *Bacillus anthracis*, a spore-forming, gram-positive bacterium that causes anthrax, an acute, infectious disease characterized by various symptoms, including hemorrhage, edema, necrosis, and more. *B. anthracis* can be found world-wide, and it has recently attracted attention because of its potential use in bioterrorism. The spores of *B. anthracis* and anthrax were considered to be the main biological risk in the world in 1996. Therefore, the development of novel antibiotics that target *B. anthracis* has become rather significant (7, 8).

The crystal structure of *B. anthracis* NADS apo-enzyme

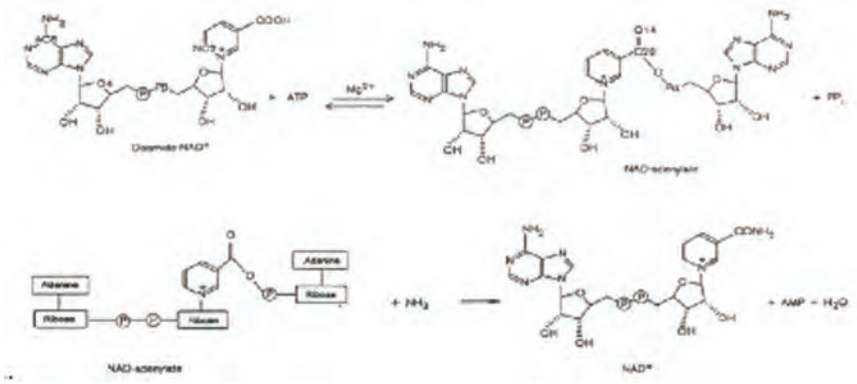

Fig. 1. The conversion of NaAD to NAD+ by NAD+ synthetase

(*ban*NADS) is displayed in Figure 2. The *B. subtilis* form of the enzyme, *bsu*NADS, has been studied more extensively and results obtained from studies of *bsu*NADS are often used to aid in understanding the structure and function of *ban*NADS. Given that *ban*NADS and *bsu*NADS have a sequence homology of at least 70%, it has been assumed that the two enzymes' three-dimensional structures are very similar: They both consist of a compact homodimer with each monomer consisting of an α and ß subunit organized into a classic Rossmann fold. The ATP binding site is located where the two subunits meet, while the NaAD binding site is located in a deep cleft lying adjacent to the α/ß switch point, which extends across the monomer-monomer interface (3, 4).

In order to obtain good crystallographic data, the protein being studied needs to be very pure. Therefore, the addition of histidine-tags to proteins is often used to make purification easier and more efficient. The His-tag facilitates protein binding to an affinity purification column charged with Ni (II). If the vector coded for an additional protease cleavage site, the His-tag may be removed and the protein purified further. This is not always done, as it is assumed that the His-tag has no effect on the protein's structure. It has been suspected that these His-tags may cause problems, such as altering the stability of the protein's three-dimensional structure or increasing the possibility of protein aggregation, but according to Carson et al., if there is any direct effect of the His-tag on the protein, it is rare (9).

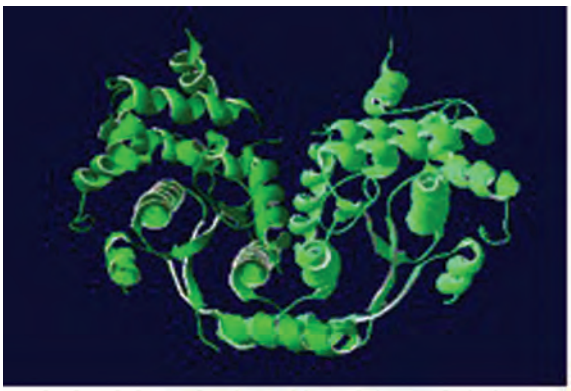

Fig. 2. Crystal structure of the B. anthracis NADS homodimer

In the case of the *ban*NADS apo-enzyme solved by McDonald et al., the C-terminal His-tag appeared to lead to some modifications of the three-dimensional structure, including the fact that residues 257–265 were missing from the apo-enzyme crystal structure but were present in *bsu*NADS and other homologous structures. This suggests that the His-tag somehow prevents these residues from adopting a stable conformation. The tag itself appeared to occupy the active sites of both monomers of *ban*NADS, but its enzymatic activity was confirmed by kinetic assay (4). It was then deemed necessary to perform an experiment that would yield structural information on the protein in solution by limited proteolysis of the *ban*NADS apo-enzyme.

Although a three-dimensional structure of a protein cannot be obtained directly from limited proteolysis, the technique has several advantages: The protein can be proteolyzed under physiological conditions, and it does not require large quantities of protein. Because proteolysis of highly ordered structures, such as α-helices and ß-strands, is thermodynamically unfavored, it is assumed that proteolysis would occur in regions of high motility and flexibility, such as loops and turns, and regions having high solvent accessibility surface areas (ASAs). This result has been confirmed in several studies, included those described by Fontana et al. (10).

Protein flexibility is vitally important to biological function, and the presence of flexible regions, such as loops and turns, aid in this function. They may also work as hinges, allowing certain domains of the protein to transport a substrate (for example) from one domain to another (11). Flexible regions also often occur between domains and may allow these domains to move about freely in solution or may aid in protein-ligand binding (12). However, these regions often do not pack into arrays well and are thus not seen in the crystal structures. Proteolysis of a protein in solution can yield information on the flexibility a protein, beyond the surface loops that are disordered in crystal structures; this technique can help determine where disordered and ordered structures exist in solution, which can then be compared to those seen in the crystal structure. In the case of *ban*NADS in particular,

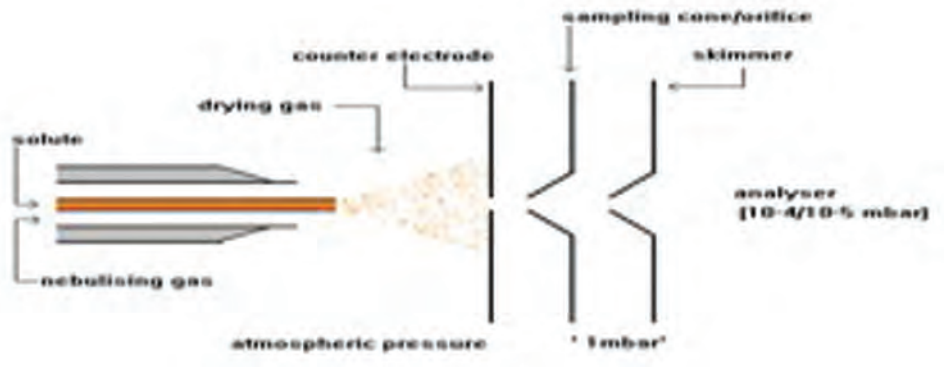

Fig. 3. Schematic of the operation of ESI-MS (Standard electrospray ionization source, Platform II) [15]

the ambiguity of the effect(s) of the C-terminal His-tag on the structure of the protein was cause for some concern: Is it in a region flexible enough to not affect enzymatic activity? Also, one would expect proteolytic cleavage to coincide with the observation of flexible regions in the structure, but if cleavage occurs in an area that appears ordered in the crystal structure, it can be suspected that this area does not exist in a rigid, ordered structure 100% of the time. Proteins are known to be dynamic in solution, so such an observation is definitely possible.

In the current experiment, the NADS apo-enzyme from *B. anthracis* is subjected to limited proteolysis by trypsin and chymotrypsin. Trypsin, which cleaves on the carboxyl side of lysine (K) and arginine (R), is a ubiquitous endoprotease found in the pancreas. Its rate of hydrolysis tends to be slower if the residue at which cleavage occurs is located beside an acidic amino acid, and proteolysis may or may not occur if a proline is present beside the cleavage site. Trypsin has a molecular weight of approximately 23.29 kDa and an optimum pH around 8.0 (13). Chymotrypsin is also a pancreatic endoprotease and cleaves on the carboxyl side of aromatic amino acids, such as tryptophan (W), tyrosine (Y), and phenylalanine (F), and sometimes leucine (L). It has a molecular weight of approximately 25 kDa and an optimum pH of 7.8 (14).

The fragments resulting from proteolysis are detected qualitatively by SDS-PAGE, and the masses of the fragments are determined quantitatively by LC-ESI-MS, or liquid chromatography electrospray ionization mass spectrometry. ESI-MS is well-suited for analyzing polar molecules which have a molecular weight between 100 Da and 1,000,000 Da (1,000 kDa). The ESI-MS itself, not including the liquid chromatography often performed in conjunction with the technique to further purify and separate fragments, operates by dissolving the sample in a polar, volatile solvent (such as ammonium acetate). The sample is then pumped through a narrow capillary tube (75–150 μM inner diameter, or i.d.) at a rate between 1 μL/min–1 mL/min. The capillary tip (not to be confused with the LC capillary) holds the sample to be ionized and sits within the ionization source of the mass spectrometer; to this tip, a strong electric voltage (3–4 kV) is applied. The resulting electric field causes the sample emerging from the tip of the capillary to be dispersed into highly-charged droplets. A nebulating gas, such as nitrogen, is used to helps direct the ionized spray to the mass spectrometer. The drying gas, also often nitrogen, flows across the ionization source and causes the solvent in the droplets to evaporate, leaving only the charged sample ions. Some of these ions pass through a sampling cone into an intermediate vacuum region and then into high vacuum region containing the analyzer, which can be one of several types, including TOF (time of flight, the type used in this experiment), quadrupole, etc. The lens voltages are set according to each individual sample (Figure 3). It is this type of mass spectrometry that is used to determine the exact masses of the fragments resulting from the proteolysis of *B. anthracis* NAD+ synthetase.

METHODS AND MATERIALS

Materials—The pET21b vector and DNase were purchased from Novagen (Madison WI), and BL21(DE3) E. coli was purchased from Invitrogen (Carisbad, CA). LB agar, LB broth, Tris, methanol, and glacial acetic acid were all purchased from Fisher Scientific. Ampicillin, all proteolytic enzymes, NH_4HCO_3, $CaCl_2$, and all other reagents were purchased from Sigma Chemical Co. (St. Louis MO).

Optical densities were measured using a Beckman DU 640B UV/Visible spectrophotometer. French press was conducted using a French Pressure Cell Press from SIM Aminco Spectronic Instruments (Rochester, NY). The HiTrap 5-mL nickel affinity chromatography column, Superdex 200 16/60 gel filtration column, Äkta Explorer purification system, and the Fractionation 900 fraction collector were purchased from Amersham Biosciences (Piscataway, NJ); the software used to run the Äkta system was Unicorn. Concentration of protein solutions and buffer exchange were conducted using an Amicon Ultracentrifugal Filter Device with a minimum molecular weight limit of 10 kDa (Millipore Corporation, Billerica, MA). SDS-PAGE was conducted using Tris-HCl 4–20% Ready Gels and Precision Plus Protein Standard from BioRad Laboratories (Hercules, CA). Coomassie Brilliant Blue was purchased from Imperial Chemical Industries. Mass spectral analyses were performed with a Tandem Q-ToF 2

Mass Spectrometer (Micromass, Manchester, UK). Liquid chromatography (for LC-MS) was performed using an LC Packings Ultimate LC Switchos microcolumn switching unit and Famos autosampler from LC Packings (San Francisco, CA). Data was analyzed with accompanying software. SDS-PAGE gels were digitized using UN-SCAN-IT gel digitizing software purchased from Silk Scientific, Inc. (Orem, UT).

Expression and Purification—A 1 L *E. coli* culture containing a pET21b vector with inserted *B. anthracis* NADS genomic DNA was grown overnight at 37°C on an LB (Luria Bertani) agar plate containing 100 μg/mL ampicillin. Multiple colonies were scraped and placed into an autoclaved 20-mL LB broth starter culture, which contained 100 μg/mL ampicillin and incubated overnight at 300 rpm and 37°C. The starter culture was then added to 1 L LB broth and incubated at 37°C, shaking at 300 rpm. Cell growth was monitored by measuring optical density (OD) at 595 nm until it reached 0.6 to 0.8. Overexpression was induced with isopropyl-beta-D-thiogalactopyranoside, or IPTG, which was added to a concentration of 1 mM, and cells were grown for an additional 2-4 hours after induction. Cells were collected by centrifugation at 6,000 g for 20 min, and the pellets were stored at -80°C.

The pellets were thawed at room temperature and dissolved in Ni Buffer A (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 10 mM imidazole, 1 mM DTT, or dithiothreitol) and protease inhibitor and DNase were added. The dissolved pellets were then lysed at 1,500 psi by French press, and the lysate was centrifuged at 1,800 rpm for 60 min and 4°C. The supernatant was separated from the pellet and sterile filtered through a 0.45-micron filter.

The filtered supernatant was purified using a nickel affinity chromatography column. The protein was run over the column at a flow rate of 2.5 mL/min in 10% Buffer B (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 500 mM imidazole, 1 mM DTT) for 10 column volumes (CV). Then, a linear gradient of Buffer B from 10 to 100% was run over 5 CV, and the gradient was held at 100% Buffer B for 3 CV. Appropriate fractions, which were determined by SDS-PAGE, were pooled and EDTA (ethylenediaminetetracetic acid) added to a concentration of 2 mM. The pooled fractions were concentrated down to a volume of approximately 3.5 mL by centrifugation.

The protein was dialyzed into gel filtration buffer (50 mM Tris, pH 7.5, 1 M NaCl, 10% glycerol, 2 mM DTT, 0.1 mM AEBSF inhibitor, or 4-(2-Aminoethyl) benzenesulfonyl fluoride). The dialyzed protein was loaded onto a Superdex 200 16/60 using a flow rate of 1 mL/min over 1.1 CV, and 2-mL fractions were collected. Protein purity was checked by SDS-PAGE.

*Limited Proteolysis of ban*NADS *with Trypsin and Chymotrypsin*—The purified NADS was buffer exchanged into 100 mM NH_4HCO_3, pH 8.5 and concentrated down to 1 mg/mL; the concentration was calculated by dividing the absorbance of the solution at 280 nm by the protein's extinction coefficient of ε = 0.728 mL/mg*cm (16) multiplied by 1 cm, according to the Beer-Lambert Law of A = εlc, where A is the absorbance of the sample, ε is the extinction coefficient of the protein, and l is the pathlength. A 1 mg/mL trypsin stock solution was prepared in 1 mM HCl, and the stock solution was diluted to 0.25 mg/mL in 100 mM NH_4HCO_3. The proteolysis contained 475 μL of 1 mg/mL NADS, 19 μL trypsin, and 54.5 μL of 100 mM $CaCl_2$, resulting in a solution containing 0.866 mg/mL NADS with the trypsin:NADS w/w ratio 1:100. The mixture was also 10 mM $CaCl_2$, and the final volume was 548.5 μL.

The tryptic proteolysis reaction was run at 25°C in a shaking incubator at 250 rpm. To stop proteolysis, a total of 80 μL of the solution was removed, 30 μL of which were frozen immediately at -80°C, while the remaining 50 μL were boiled in 16.5 μL 4X SDS-PAGE loading buffer (250 mM Tris-HCl, pH 6.8, 6% SDS, 300 mM DTT, 30% glycerol, and 0.02% bromophenol blue) at 100°C and then stored at -20°C. This was done at 0, 15, 30, 60, 120, and 180 min after proteolysis was initiated.

A similar procedure was followed for the limited proteolysis using chymotrypsin. The *ban*NADS was buffer exchanged into 100 mM Tris-HCl, pH 7.8. The chymotrypsin stock solution consisted of 1 mg/mL chymotrypsin in 1 mM HCl and 2 mM $CaCl_2$. Instead of a w/w ratio of 1:100 (enzyme:protein), a ratio of 1:60 was used, as recommended by Sigma. The chymotrypsin stock solution was diluted to 0.4 mg/mL, and 19.5 μL was added to 475 μL NADS and 54.5 μL 100 mM $CaCl_2$. The final volume was 549 μL, and the solution was 0.865 mg/mL in NADS, 10 mM $CaCl_2$ and 0.0138 mg/mL in chymotrypsin. The reaction was run at 37°C in a shaking incubator at 250 rpm. Proteolysis was halted in the same manner as before. Another chymotryptic proteolysis was conducted with the w/w ratio of 1:25.

Determination of Fragment Molecular Weight—The detection of proteolytic fragments was determined qualitatively by SDS-PAGE. The 6 time points taken during each reaction were thawed, and 20 μL of each sample were loaded onto a 4-20% polyacrylamide gel; 10 μL of the protein standard were loaded. The gels were run at 100 V, after which they were stained with Coomassie Blue (30% methanol, 12% acetic acid, 0.1% Commassie Brilliant Blue) for 1 hr on a shaker. The gels were then de-stained (30% methanol, 12% acetic acid) overnight and scanned and edited for contrast improvement using Adobe Photoshop.

Molecular weights from the gel were calculated by determining the rf (retention factor) values by gel digitization using UN-SCAN-IT software. The rf values of the protein standard were plotted against the logMW of the protein standard to yield a linear graph with an equation of the form y = mx + b. The resulting equation was used to calculate the molecular weights of the proteolytic fragments.

After analyzing the SDS-PAGE gels, four samples were selected for analysis by electrospray LC-MS: For the tryptic digestion, the t = 60 min and t = 120 min samples were selected, and for both the 1:25 and 1:60 w/w enzyme:protein chymotryptic digestions, the t = 120 min samples were analyzed. The samples were concentrated on a 300 μm i.d. (inner

Table I:
Proteolytic Fragment Masses Determined by SDS-PAGE and LC-MS

LC-MS				SDS-PAGE		
Tryptic Digest 1:100 (t=60 min)				**Tryptic Digest 1:100**		
Mass	*%Error*	*Fragment*		*Mass*	*Dev*	*Fragment*
31493.34	2.35×10^{-3}	intact		34134.24	±1611.42*	intact
25064.31	9.18×10^{-3}	A10-K239		27205.11	±188.92	A10-K239
22015.55	1.59×10^{-4}	D87-H284		24040.10	±735.61	D87-H284
16693.57	2.22×10^{-4}	D87-K239		16966.42	0.00	D87-K239
9495.13	9.53×10^{-5}	M1-K86		10372.16	±124.50	M1-K86
3684.99	*2.86×10^{-4}*	*M207-K239*		29709.00	±413.44	A10-H284
4185.48	*2.54×10^{-4}*	*Q96-K132*		8873.19	0.00	--
s=	**3.35×10^{-3}**					
Tryptic Digest 1:100 (t=120 min)						
31493.34	2.35×10^{-3}	intact				
22015.55	1.59×10^{-4}	D87-H284				
3684.99	*2.86×10^{-4}*	*M207-K239*				
s=	**1.23×10^{-3}**					
Chymotryptic Digest 1:60 (t=120 min)				**Chymotryptic Digest 1:60**		
31494	4.45×10^{-3}	intact		34450.92	±655.13	intact
30540.70	3.18×10^{-4}	M1-L277		78723.51	±541.50	--
20974.77	2.22×10^{-4}	H12-Y204		24525.38	±353.36	M1-Y229
29972.97	2.22×10^{-4}	M1-W271		13294.00	±394.03	--
s=	**2.10×10^{-3}**			10731.31	±427.92	M1-F97
Chymotryptic Digest 1:25 (t=120 min)						
31494.00	4.45×10^{-3}	intact				
30540.70	3.18×10^{-4}	M1-L277				
20974.77	2.22×10^{-4}	H12-Y204				
20837.63	9.53×10^{-5}	V13-Y204				
s=	**2.12×10^{-3}**					

Red content was not observed on SDS-PAGE gel (no corresponding mass seen in Lane 10).
Blue content had no corresponding mass from LC-MS
*Standard deviations of fragments observed on SDS-PAGE were calculated according to the equation:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

where σ is the standard deviation, N is the number of masses corresponding to each fragment, xi is the ith fragment mass, and $\bar{x}$ is the average of the masses corresponding to the fragment in question.

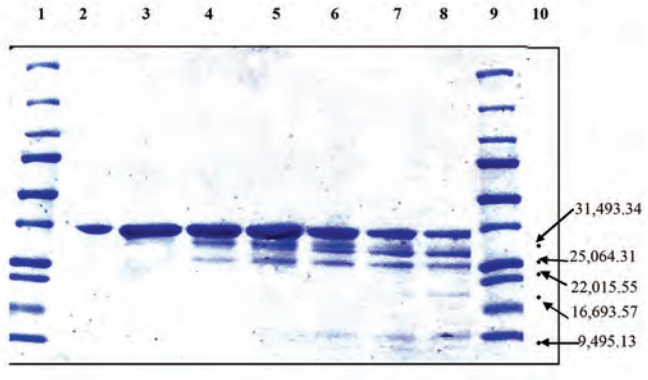

Fig. 4a. Limited Proteolysis of *ban*NADS with Trypsin (1:100). Lane 1) Std, 2) *ban*NADS (1 mg/mL), 3) t = 0 min, 4) t = 15 min, 5) t = 30 min, 6) t = 60 min, 7) t = 120 min, 8) t = 180 min, 9) Std, 10) retention factors calculated from masses obtained by LC-MS

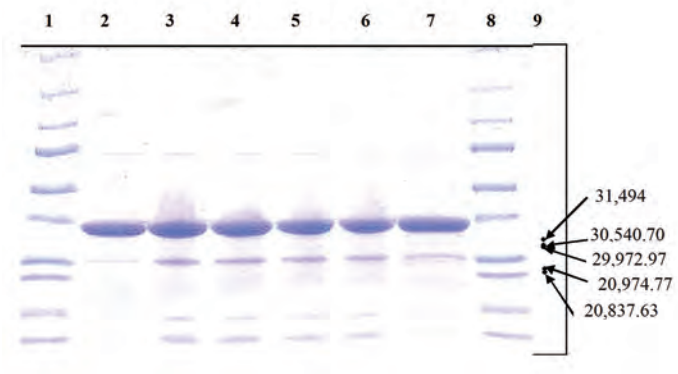

Fig. 4b. Limited Proteolysis of *ban*NADS with Chymotrypsin (1:60). Lane 1) Std, 2) t = 0 min, 3) t = 15 min, 4) t = 30 min, 5) t = 60 min, 6) t = 120 min, 7) t = 180 min, 8) Std, 9) retention factors calculated from massed obtained by LC-MS

diameter) C18 precolumn at a flow rate of 10 µL/min with 0.1% formic acid and then flushed onto a 75 µm i.d. C18 column at 200 µL/min with a gradient of 5–100% acetonitrile (0.1% formic acid) in 30 min. The nano-LC interface was used to transfer the LC eluent into the mass spectrometer, where fragments were detected and their masses determined. The resulting data was analyzed by the software that accompanied the mass spectrometer, and the sequences of the fragments were determined.

The PDB file of the *ban*NADS apo-enzyme was viewed using Swiss-PdbViewer software (17), and solvent accessible surface areas and B-factors were calculated using Gerstein's Calc-surface program (18).

RESULTS

SDS-PAGE—The molecular weight of the proteolytic fragments were calculated from the rf values of the SDS-PAGE gels, which are displayed in Figure 4. The molecular weights of the fragments that were determined by LC-MS were plotted alongside the gels for comparison.

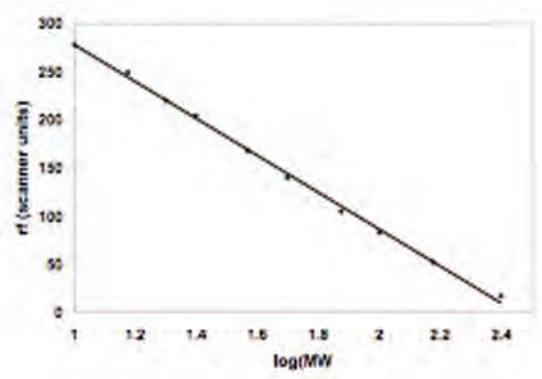

Fig.5. Rf (protein standard) vs. log(MW). The rf values of the protein standard were plotted against log(MW) to generate an equation of the form y = mx + b, which was used to calculate the molecular weights of the fragments resulting from the tryptic proteolysis. 1 scanner unit = 0.01 in.

The rf values of the protein standard were plotted against the logMW (of the standard) to generate the linear graph; the graph used to calculate molecular weights from the tryptic digestion is displayed in Figure 5. A similar graph was generated for the SDS-PAGE gel run for the chymotryptic proteolysis.

The graphs produced for the gel corresponding to the proteolysis conducted with trypsin and chymotrypsin yielded Equations 1 and 2, respectively:

$$rf = -191.82\log(MW) + 469.86 \quad (1)$$

$$rf = -182.53\log(MW) + 457.24 \quad (2)$$

The R2(1) was 0.9981, and the R2(2) was 0.9971.

LC-MS—Even though the masses of the fragments could be calculated from the SDS-PAGE, this method is usually believed to be rather inaccurate due to the tendency of some proteins (such as *ban*NADS) to have higher molecular weights according to SDS-PAGE than the known molecular weights, so the masses were also determined by electrospray LC-MS. Sample mass spectra are displayed in Figure 6. The results from the mass spectrometry, along with the corresponding peptide sequences, and the masses resulting from SDS-PAGE are reported in Table I.

Red content was not observed on SDS-PAGE gel (no corresponding mass seen in Lane 10).

Blue content had no corresponding mass from LC-MS

*Standard deviations of fragments observed on SDS-PAGE were calculated according to the equation: where α is the standard deviation, N is the number of masses corresponding to each fragment, xi is the ith fragment mass, and $\bar{x}$ is the average of the masses corresponding to the fragment in question.

Although SDS-PAGE confirmed the presence of proteolytic fragments, especially the gel displaying samples from

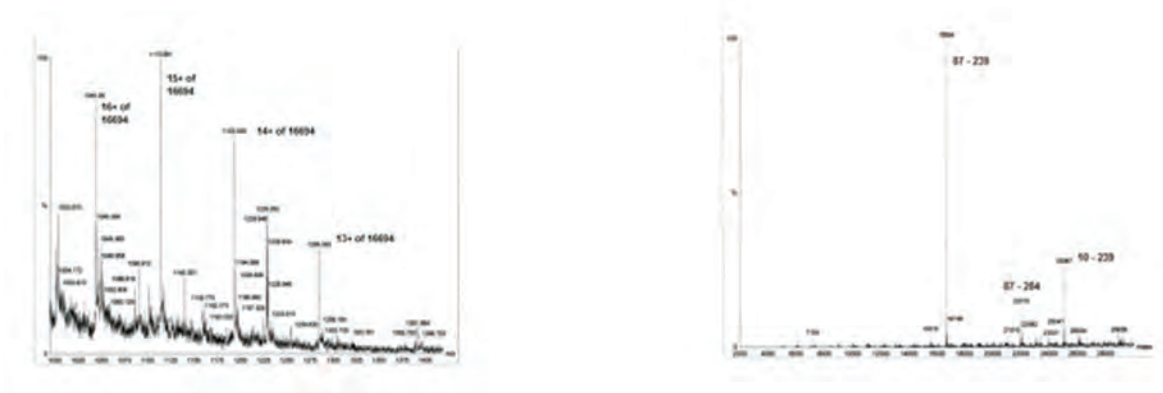

Fig. 6. Sample mass spectra resulting from LC-MS.
a) Raw data resulting from LC-MS of sample of tryptic proteolysis at t = 60 min;
b) Resulting molecular masses after analysis of raw data for the sample used in 6a

tryptic proteolysis, LC-MS generated much more accurate and precise mass values than did the SDS-PAGE in terms of actual mass values, with both percent errors and standard deviations of less than 0.1%. Indeed, the gel displayed in Figure 4a accurately depicts the presence of tryptic fragments, but only qualitatively. While the majority of fragment masses resulting from the tryptic digest (as determined by LC-MS) appeared to have corresponding bands on the SDS-PAGE gel, the same could not be assumed for the results of the chymotryptic proteolysis (Fig. 4). In fact, most of the bands seen on the gel run on the chymotrypsin proteolysis samples were not seen in the mass spectra, and no reasonable cleavage sites could be found to correspond to two of the bands seen on the gel. There were also two bands on the gel that was run on the trypsin proteolysis samples that were not seen in the mass spectra, one of which could not be matched with a conclusive cleavage pattern.

The observed cleavage sites are displayed on the primary structure sequence map of *ban*NADS in Figure 7. There were some fragments that were expected to be seen in the mass spectra but were not. These were determined by observing flexible and exposed regions on the crystal structure of *ban*NADS and are displayed on the crystal structure in Figures 8 and 9, respectively.

MTLQEQIMKA10 LHVQPVIDPK20 AEIRKRVDFL30 KDYVKKTGAK40

GFVLGISGGQ50 DSTLAGRLAQ60 LAVEEIRNEG70 GNATFIAVRL80

PYKVQKDEDD90 AQLALQFIQA100 DQSVAFDIAS110 TVDAFSNQYE120

NLLDESLTDF130 NKGNVKARIR140 MVTQYAIGGQ150 KGLLVIGTDH160

AAEAVTGFFT170 KFGDGGADLL180 PLTGLTKRQG190

RALLQELGAD200 ERLYLKMPTA210 DLLDEKPGQA220 DETELGITYD230

QLDDYLEGKT240 VPADVAEKIE250 KRYTVSEHKR260

QVPASMFDDW270 WKLAAALEHH280 HHHH

Most of the tryptic cleavages occurred on the carboxyl side of lysine; although, there was one anomalous cleavage at Leu-95. The only anomalous chymotryptic cleavage occurred at His-12, but the concentration of this fragment (Val-13–Tyr-204) appeared to be relatively small, according to the intensity of the corresponding peak seen in the mass spectra (not shown). One cleavage occurred in the P1 loop (residues 85-87), and two occurred in the P2 loop (206–226). Two cleavages (both by chymotrypsin) occurred within the C-terminal α-helix (266–284). There were six cleavages by trypsin and five cleavages by chymotrypsin.

Location of Cleavage Sites—The three-dimensional structure of the *ban*NADS apo-enzyme was viewed using Swiss-PdbViewer, and Figures 8 and 9 were generated. Lys-9 appeared to be located at the end of an α-helix, as did Lys-132. Leu-11 and His-12 were located in the loop connected

Table II:
Solvent Accessibility Surface Areas of
Observed and Expected Sites of Cleavage

Observed			Expected		
Residue	$ASA_{N\ or\ C}$ ($Å^2$)	$ASA_{residue}$ ($Å^2$)	*Residue*	$ASA_{N\ or\ C}$ ($Å^2$)	$ASA_{residue}$ ($Å^2$)
K9**	**0.00	141.72	**Y82**	0.00	44.28
A10***	0.24	51.22	K83	0.00	115.48
	ASA_{dimer}	***192.94***		***ASA_{dimer}***	***159.76***
L11	0.00	15.47	**K83**	0.08	115.48
H12	0.00	121.28	V84	0.44	150.68
	ASA_{dimer}	***136.75***		***ASA_{dimer}***	***266.16***
H12	1.44	121.28	**F97**	0.00	29.42
V13	0.00	8.71	I98	0.00	0.05
	ASA_{dimer}	***129.99***		***ASA_{dimer}***	***29.47***
L95	1.18	24.25	**Y119**	0.00	0.00
Q96	0.00	112.98	E120	0.00	67.33
	ASA_{dimer}	***137.23***		***ASA_{dimer}***	***67.33***
K132	0.00	49.54	**R202**	0.00	147.69
G133	0.00	8.06	L203	0.00	7.36
	ASA_{dimer}	***57.60***		***ASA_{dimer}***	***155.05***
K239	0.99	91.78	**Y229**	0.27	69.21
T240	5.63	139.23	D230	0.56	77.68
	ASA_{dimer}	***231.01***		***ASA_{dimer}***	***146.89***
W271	0.00	26.11	**R252**	0.00	82.55
K272	0.00	27.56	Y253	0.00	26.48
	ASA_{dimer}	***53.67***		***ASA_{dimer}***	***109.03***
L277	1.19	26.06	**Y253**	0.00	26.48
E278	0.61	83.50	T254	0.00	94.01
	ASA_{dimer}	***109.56***		***ASA_{dimer}***	***120.49***
			F267	0.00	66.36
			D268	0.00	0.00
				ASA_{dimer}	66.36
			K272	1.42	27.56
			L273	0.30	25.87
				ASA_{dimer}	***53.43***

*Residues in bold were located on the N-terminal side of the peptide bond. **The value reported beside the bolded residues is the ASA of $C_{carbonyl}$. ***Non-bolded residues were located on the C-terminal side of the peptide bond. The ASA reported beside it is of this residue's amide nitrogen. *Italicized* numbers are the sum of two consecutive residues involved in peptide linkage.

to Lys-9, while Lys-239 was found in another nearby loop. Lys-86 was missing from the structure due to the lack of packing of the P1 loop, as were Tyr-204 and Lys-206 from the P2 loop. Leu-95 was located, unexpectedly, in the middle of an α-helix, but the α-helix appeared to be highly exposed to solvent. Similarly, Trp-271 and Leu-277 were located within the C-terminal helix containing the His-tag. Sites where cleavage was expected to occur but did not were determined by inspection; they were most often located at the ends of helices or in flexible surface loops.

In order to gain more insight into why there was no cleavage at the sites highlighted in Figure 9, the solvent accessible surface areas of the peptidyl nitrogen and $C_{carbonyl}$ atoms involved in peptide linkage and of the residues in their entirety were calculated using Gerstein's Calc-surface program, which uses a probe size of 1.4 Å, and the results are reported in Table II.

The magnitude of solvent accessibility was determined by comparing the ASA values reported in Table II with the values in Table III. They are the sum of the $ASA_{nonpolar}$ and ASA_{polar} given by the Richards method (19). These ASA values are assumed to be a measure of the accessibility of the unbound residues (that is, residues not bound in a protein).

Table III:
ASA Values of Ala-X-Ala Tripeptides

*Ala-X-Ala**	ASA_{total} ($Å^2$)
F	214.4
H	187.6
K	217.4
L	186.0
R	250.0
W	255.6
Y	231.2

*X is the amino acid residue.

Notice that the values in Table III are significantly higher than the values reported in Table II. This was expected because amino acid residues surrounded by other structural components of a protein are certainly less exposed to solvent than are residues in an unbound tripeptide. More information on the flexibility of the atoms involved in cleavage was obtained by determining the B-factors of the $C_{carbonyl}$ (of the residue at which cleavage occurred on the carboxyl side) and the peptidyl N. These values are reported in Table IV. Because the B-factor is a measure of flexibility and often reflects the uncertainty of the position of the atom in the crystal structure (20), the residues where cleavage was expected appear, in general, to be more mobile and flexible than the sites where cleavage was actually observed.

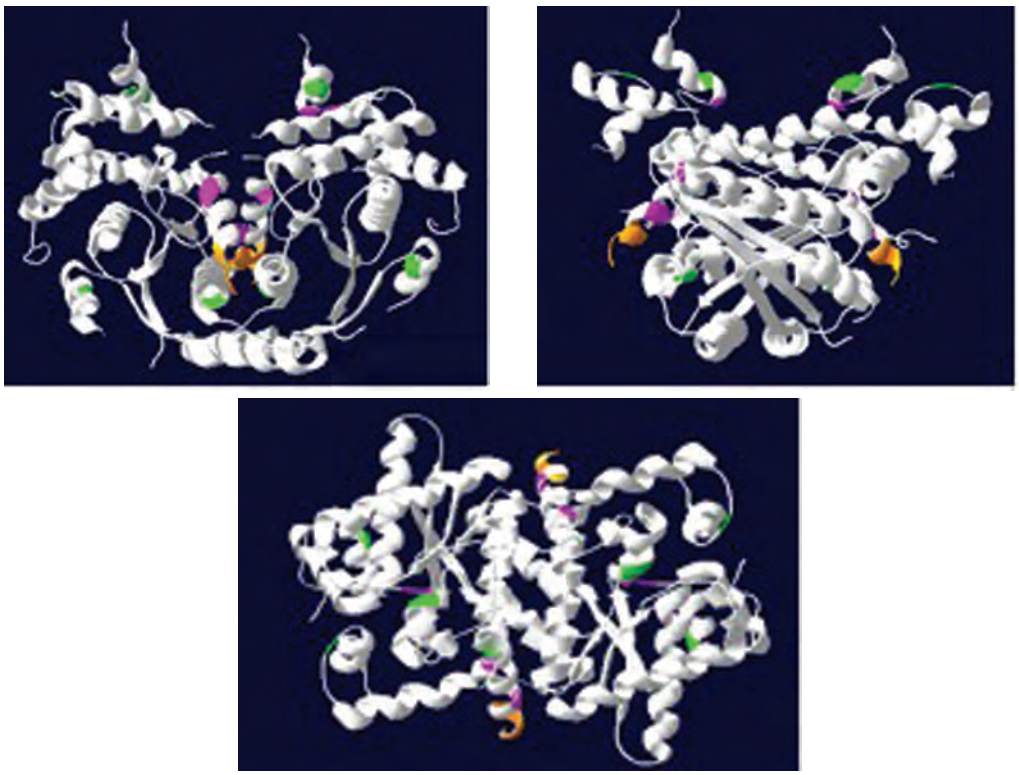

Fig. 8. Observed proteolytic cleavages. Crystal structure of *ban*NADS containing the His-tag viewed from three different angles. Sites in green (K9, L95, K132, and K239) were cleaved by trypsin according to LC-MS. K86 and K206 are not seen due to lack of electron density in the P1 and P2 loops, respectively. Sites in magenta (L11, H12, W271, and L277) were cleaved by chymotrypsin according to LC-MS. The His-tag is highlighted in orange.

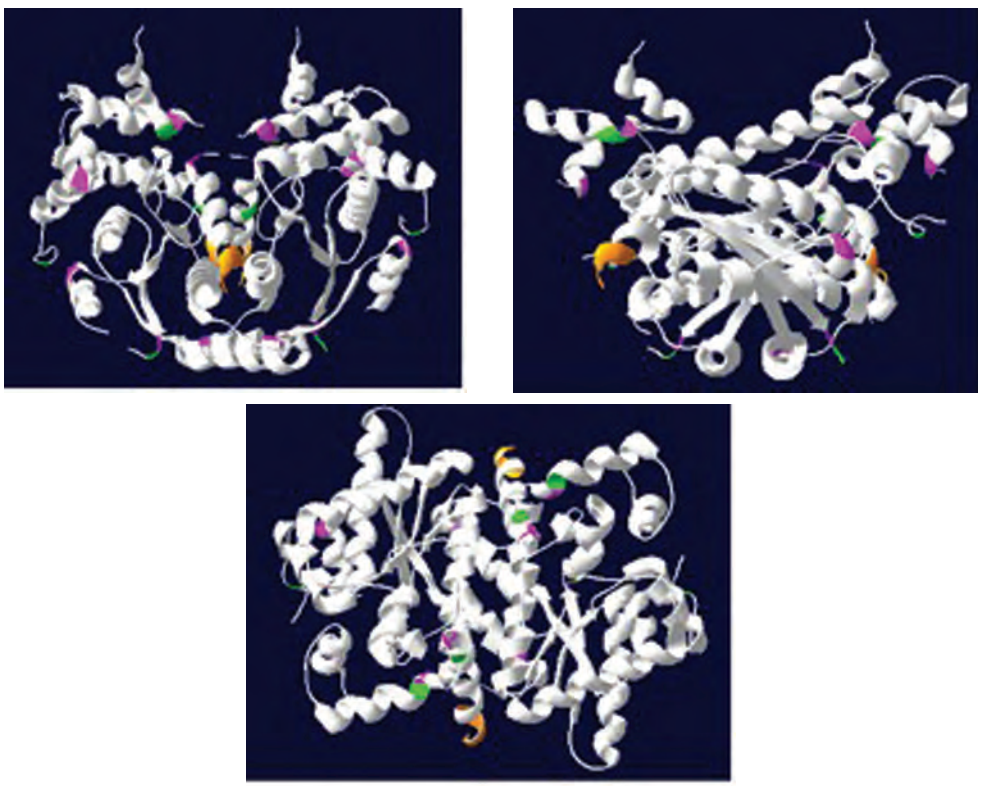

Fig. 9. Expected proteolytic cleavages. Crystal structure of *ban*NADS containing the His-tag viewed from three different erent angles. Expected tryptic cleavages (K83, R202, R252, and K272) are shown in green. Expected chymotryptic cleavages (Y82, F97, Y119, Y229, Y253, and F267) are shown in magenta. The His-tag is highlighted in orange.

Table IV
B-Factors of Observed and Expected Sites of Cleavage

Observed		Expected	
Residue	*B-Factor$_{C\ or\ N}$ (Å^2)*	*Residue*	*B-Factor$_{C\ or\ N}$ (Å^2)*
***K9**	**38.08	**Y82**	31.82
***A10	37.05	K83	35.21
L11	34.43	**K83**	44.78
H12	33.43	V84	48.51
H12	32.71	**F97**	27.74
V13	29.56	I98	26.52
L95	28.79	**Y119**	25.43
Q96	29.96	E120	24.25
K132	23.75	**R202**	47.97
G133	23.84	L203	47.72
K239	35.04	**Y229**	43.94
T240	36.00	D230	43.36
W271	28.00	**R252**	59.72
K272	27.31	Y253	58.72
L277	31.79	**Y253**	60.17
E278	32.11	T254	61.58
		F267	48.31
		D268	44.66
		K272	25.78
		L273	26.13

*Residues in bold were located on the N-terminal side of the peptide bond.
**The value reported beside the bolded residues is the B-Factor of $C_{carbonyl}$.
***Non-bolded residues were located on the C-terminal side of the peptide bond. The B-Factor reported beside it is of this residue's amide nitrogen.

DISCUSSION AND CONCLUSIONS

Even though protein structures solved by x-ray diffraction are generally considered to be similar due to the high percentage of solvent (i.e., water) present in a crystal of a protein, the technique has its disadvantages. One of the main drawbacks is that, in order to solve any structure using x-ray crystallography, the molecule(s) being studied must be in the solid state. However, the native state of a protein is dynamic and is found in solution, and it is this form of the protein in which we are interested (11).

Because there was some evidence suggesting that the His-tag on the *ban*NADS apo-enzyme may have affected the protein's crystal structure, limited proteolysis of the protein seemed appropriate in order to gain more information on the native protein (that is, the protein found in solution). The NADS was subjected to proteolysis by trypsin and chymotrypsin, and the progression of the proteolysis was observed by SDS-PAGE. The tryptic proteolysis samples were run on the gel displayed in Figure 4a. The gel shows quite nicely how the concentration of *ban*NADS decreased over time, while the amount of proteolytic fragments increased. The fragment masses determined by LC-MS appeared to correspond well to several of the fragments seen on the gel. In contrast, in Figure 4b, the concentration of NADS does not appear to decrease over time, and the fragment masses determined by mass spectrometry do not appear to correlate well with the bands seen on the gel. Because this was not the case with the gel displaying the results of the proteolysis with trypsin, it can be assumed that there was a problem with the samples run on the gel displayed in Figure 4b, and/or there were experimental complications concerning the way in which the gel was run.

SDS-PAGE Provides Information on the Cleavage Process—Interestingly, a fragment seen in Figure 4a, which has a calculated molecular weight of approximately 30 kDa, seems to indicate a possible single cleavage at Lys-9, resulting in a fragment containing residues 10–284 (and another containing residues 1–9). The size of this band appears to decrease between 15 and 180 min into the reaction, but the band corresponding to the fragment containing residues 10–239 (~25 kDa) (Table I) seems to increase in size. This information is consistent with the idea that the fragment containing residues 10–239 resulted from two consecutive cleavages: 1) The protein was first cleaved at Lys-9, resulting in the large fragment (10–284), and 2) this large fragment was cleaved again at Lys-239. Similarly, LC-MS detected fragments corresponding to residues 1–86 (~9.5 kDa), 87–284 (~22 kDa), and 87–239 (~16.7 kDa) (Table I). All three of these fragments also appear on the gel (Fig. 4a), with the concentrations of fragment 1–86 and fragment 87–239 increasing over time while the concentration of fragment 87–284 appeared to remain relatively constant. Because fragment 87–239 did not appear until 120 min, and fragment 87–284 first appeared after only 15 min, one is led to assume that there was first a cleavage at Lys-86 and then another at Lys-239, resulting in the fragment containing residues 87–239. The reason for this sequence of cleavages is unclear; perhaps it is a result of the short length of the loop containing Lys-239, which extends only from residues 237–242, and its location between α-helices 11 and 12 (4). Another fragment (not seen on the SDS-PAGE gel) corresponds to residues 207–239; the results, however, give no definite explanation for how this fragment was produced, but this fragment does further suggest that Lys-239 is readily cleaved by trypsin. One possible explanation for this cleavage pattern could be that, when cleavages that disrupt the homodimer occur, the protein unfolds, allowing for more secondary cleavages.

SDS-PAGE of Chymotryptic Proteolysis Samples—Because the gel displayed in Figure 4b is probably not a good representation of the proteolytic reactions that occurred during the experiment, it is difficult to determine the means by which fragment 12–204 and fragment 13–204 were generated. There are also no data from mass spectrometry supporting the assumption of sequential cleavage to produce these fragments. Fortunately, there were a couple fragments resulting from

single cleavages detected by mass spectrometry, including those containing amino acids 1–271 and 1–277. The bands seen on the gel, not including the one corresponding to banNADS, may be impurities in the sample, but it seems unlikely that they are resultant of proteolysis with chymotrypsin.

Observed vs. Expected Sites of Cleavage—Several of the cleavage sites discussed above (Lys-9, Lys-86, Lys-206, and Lys-239) coincided well with the crystal structure of *ban*NADS; that is, these sites appeared to be located in or contiguous with seemingly flexible regions, such as loops found on the surface of the protein. However, there was definitely one cleavage that was not expected: Leu-95. Not only does trypsin cleave almost always on the carboxyl side of lysine and arginine, but this residue appeared to be located in the middle of helix α4 (4), even though this helix appears to be very solvent-exposed. Fragment 96–132 is included in the results, though, because of the relatively large intensity of the mass spectral peak (not shown). Despite the significance of this anomalous result, an explanation of why tryptic proteolysis occurred at Leu-95 is lacking. Perhaps the relatively large accessible surface area (ASA) of 137.23 Å2 (Table II) contributes to promotion of cleavage at this site.

The ASA of both observed sites of cleavages and expected sites of cleavage were calculated, and from the data reported in Table II, it seems that, in general, the observed cleavages occurred in areas having higher ASA values. Because the PDB file of *ban*NADS lacked residues 85–87 (P1 loop), 206–226 (P2 loop), and 257–265 and because sites of expected cleavage were determined only by inspection, there are some data "missing." That is, several of the latter sites listed, such as residues 82, 83, 202, 229, 252, etc., cannot really be compared to actual similar sites where cleavage was observed in terms of ASA because ASA values of sites of observed cleavages could not be determined. It is quite possible, then, that these observed cleavage sites had even higher ASA values than the ones listed in Table II.

The chymotryptic cleavages at 271 and 277 also occurred within an α-helix, but cleavage of peptide bonds in highly-ordered, rigid secondary structures is thermodynamically unfavored. This is because, in order for cleavage to occur in a helix, stabilizing interactions, such as hydrogen bonds, would have to be broken. Therefore, if these residues were located in α-helices 100% of the time that the protein was in solution, then cleavage by limited proteolysis would not occur. The results, then, suggest that these regions of secondary structure are not always locked into rigid helices, but rather, they are dynamic and flexible, at least sometimes. Indeed, McDonald et al determined that residues 266–284 were found in the form of a helix in the apo-enzyme. When ligand was bound to the crystallized protein, the C-terminal helix became a more flexible random coil (4).

The results of the limited proteolysis then support the hypothesis that the structure of *ban*NADS in solution differs from that of the crystallized form. Perhaps this is a direct result of the presence of the His-tag, which may, through its interaction with other components of the protein, such as its contacts with α12, α6, and α7 (4), have caused structural changes resultant of the destabilization of some areas of the protein (and stabilization of other areas) and/or abnormal interactions among the amino acid residues.

Future Experiments—Even though much information was obtained from this experiment, some of the results are inconclusive, and more experiments need to be performed in order to make more concrete conclusions. Suggested experiments to perform in addition to the one discussed in this paper included running the proteolysis using another protease such as elastase or V8 protease. Limited proteolysis with chymotrypsin should probably be repeated to obtain a gel that more accurately displays the results of chymotryptic cleavage. Other techniques, such as CD-spectroscopy, can also be used to complement the experimental results from limited proteolysis.

ACKNOWLEDGMENTS

This experiment was funded by a grant from the NIH and would not have been possible without the invaluable assistance and advice of my Honors Committee: Drs. Christie Brouillette, David Graves, and Aaron Lucius (and Dr. Gary Gray, Honors adviser). Special thanks also to Marion Kirk (UAB Comprehensive Cancer Center), who conducted the LC-MS, and to Dr. Pamela Pruett, Dr. Irina Protassevitch, Dr. Heather McDonald, Tasha Kane, and Wendy Yang for helping me during my time spent at UAB CBSE.

REFERENCES

1. Rizzi, M., Nessi, C., Mattevi, A., Coda, A., Bolognesi, M., and Galizzi, A. (1996) EMBO.15, 5125-5134
2. Rizzi, M. and Schindelin, H. (2002) Curr. Opin. Struct. Biol. 12, 709-720
3. Deredjiev, Y., Symersky, J., Singh, R., Jedrzejas, M., Brouillette, C., Brouillette, W., Muccio, D., Chattopadbyay, D., and DeLucas, L. (2001) Acta Cryst. D57, 806-812
4. McDonald, H. M., Pruett, P., Deivanayagam, C., Protassevitch, I., Carson, S. M., DeLucas, L. J., Brouillette, W. J., and Brouillette, C. G. (2007) Submitted
5. Rizzi, M., Bolognesi, M., and Coda, A. (1998) Structure 6, 1129-1140
6. McDevitt, D., Payne, D., Holmes, D., and Rosenberg, M. (2002) J. Appl. Microbiol. 92, 28S-34S
7. Center for Disease Control and Prevention. Department of Health and Human Services. Center for Infectious Diseases / Division of Bacterial and Mycotic Diseases (2005) [Online]. Available from http://www.bt.cdc.gov/ Accessed 2006 July 10
8. Jedrzejas, M. J. (2002) Crit. Rev. Biochem. Mol. Biol. 37, 339-373
9. Carson, M., Johnson, D., McDonald, H., Brouillette, C., and DeLucas, L. (2007) Acta Cryst. D63, 295-301
10. Fontana, A., Polverino de Laureto, P., Spolaore, B., Frare,

E., Picotti, P., and Zambonin, M. (2004) Acta Biochem. Polonica 51, 299-321
11. Carlson, G. (ed.) (2006) Principles of Biochemistry. (4th Ed.) New Jersey: Prentice Hall
12. [Online]. Available from http://molpharm.aspetjounrals.org/ Accessed 2007 April 12
13. Sigma-Aldrich, Inc. Proteomics Guide: Trypsin
14. Sigma-Aldrich, Inc. Enzyme Explorer: Chymotrypsin
15. Ashcroft, A. E. An Introduction to Mass Spectrometry. (2007) [Online]. Available from http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm Accessed 2007 February 1
16. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server: (in) John M. Walker (ed.): The Proteomics Protocols Handbook, Humana Press (2005) pp. 571-607. [Online]. Available from http://www.expasy.ch/tools/protparam.html Accessed 2006 June 14
17. Gues, N. and Peitsch, M. C. (1997) Electrophoresis 18, 2714-2723
18. Gerstein, M. (1992) Acta Cryst. A48, 271-276. [Online]. Available from http://molbio.info.nih.gov/structbio/basic.html Accessed 2007 April 12
19. Baker, B. and Murphy, M. (1998) Methods in Enzymology 295, 294-315
20. Radivoiac, P., Obradovic, Z., Smith, D. K., Zhu, G., Vucetic, S., Brown, C. J., Lawson, J. D., Dunker, A. K. (2004) Protein Science 13, 71-80