Mass Instability in Isolated Recombinant FixL Heme Domains of *Bradyrhizobium japonicum*†

James D. Satterlee,*,‡ Christine Suquet,‡ Anil K. Bidwai,§,|| James E. Erman,§ Linda Schwall,⊥ and Ralph Jimenez⊥

Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, Department of Chemistry and Biochemistry, Northern Illinois University, DeKalb, Illinois 60115, and JILA, University of Colorado and National Institute of Standards and Technology, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309

Received August 17, 2007; Revised Manuscript Received December 6, 2007

ABSTRACT: Several recombinant *Bradyrhizobium japonicum* FixL heme domains (*Bj*FixLH) have been characterized and their temporal mass stabilities assessed by MALDI-TOF mass spectrometry. The intact heme domains all bound heme and gave normal UV−visible spectra, indicating that they were correctly assembled. Proteins produced at Washington State University included a parent 131-amino acid "full-length heme domain" (FLHD) of primary sequence T_{140}−Q_{270} (*Bj*FixLH$_{140-270}$), a histidine-tagged analogue containing an N-terminal extension, and five different terminus-truncated variants. The smallest of these was a 106-amino acid "core PAS heme domain" with primary sequence T_{151}−L_{256}. All variants except for the smallest exhibited significant mass instability, assessed by MALDI-TOF mass spectrometry, that was apparent within 1−16 days standing in a sterile environment at room temperature. Two full-length heme domains expressed independently in geographically remote laboratories (Northern Illinois University and JILA, University of Colorado) also exhibited this mass instability. A mass loss of as much as ∼25% of the starting mass has been observed, which could explain the "missing" terminal amino acids in published crystal structures. This work documents the phenomenon and its persistence despite (i) sample sterilization, (ii) protease inhibitors, (iii) primary sequence variations, (iv) the presence or absence of ferriheme ligands, and (v) the presence or absence of O_2.

The two-domain, heme sensing/signaling protein FixL and its response regulator FixJ are critical upstream regulatory proteins for symbiotic bacterial nitrogen fixation (*1*−*8*). In the case of *Sinorhizobium meliloti* (*Sm*),[1] formerly *Rhizobium meliloti* (*Rm*), the requirement for *Sm*FixL is absolute (*2*, *4*, *7*). Without *Sm*FixL, nitrogen fixation would not occur. In the case of *Bradyrhizobium japonicum* (*Bj*), the requirement for BjFixL is less absolute, as it is one of two recognized regulators of the cascade activation that produce the complement of proteins required for nitrogen fixation (*5*, *8*).

The role of FixL in both symbiotic bacteria is to sense the ambient oxygen level and to signal for nitrogen fixation cascade regulation. Whereas the details have been more completely elucidated for *S. meliloti* (*2*−*5*, *7*, *9*−*11*), the *Bj*FixL-related process is believed to be similar (*5*, *6*, *8*). FixL from both species consists of two distinct domains, one an O_2-sensing domain that contains heme and the other a histidine kinase domain that catalyzes FixJ phosphorylation and activation (*9*−*12*). In aerobic environments when the FixL heme is saturated with molecular oxygen, FixL does not catalytically phosphorylate transcription factor FixJ. Unphosphorylated FixJ is inactive, and nitrogenase is therefore not made. However, at a low O_2 partial pressure, formation of deoxy-FixL results in phosphorylation of FixJ via ATP, which, in turn, activates transcription of *nif* and *fix* genes ultimately resulting in nitrogen fixation (*2*−*12*).

Proper FixL function requires that the oxygenation status of the PAS−heme−Fe^{2+} domain be conveyed to the kinase domain. This idea led to the hypothesis of an interdomain signal (*13*, *14*) that transmits heme oxygenation status in the sensing domain to the kinase domain. Such an interdomain activation signal must have as its initial component an intradomain signal, which is initiated by dissociation of O_2 from the heme−Fe^{2+}−O_2 complex in the sensing domain and subsequently propagates within the heme domain before passing to the kinase domain.

The *Sm*- and *Bj*FixL heme domains are attractive research targets (*10*−*13*, *15*−*22*) because they are comparatively small (∼15 kDa), soluble globular proteins in which the structure and dynamics of the initial steps in chemical signal

† This work was supported in part by NIH Grant RO1GM47645 (J.D.S.), NIH Grant R15 GM59740 (J.E.E.), and the National Institute of Standards and Technology (R.J.).

* To whom correspondence should be addressed. Phone: (509) 335-6681. Fax: (509) 335-8867. E-mail: hemeteam@wsu.edu.

‡ Washington State University.

§ Northern Illinois University.

|| Current address: Department of Molecular and Cellular Biochemistry, The Ohio State University, Columbus, OH 43220.

⊥ University of Colorado and National Institute of Standards and Technology.

[1] Abbreviations: *Bj*, *Bradyrhizobium japonicum*; *Sm*, *Sinorhizobium meliloti*; *Bj*FixLH, generic reference to *B. japonicum* FixL heme domains; *Bj*FixLH$_{140-270}$, WSU and NIU full-length recombinant PAS heme domain of *Bj*FixL; EDTA, ethylenediaminetetraacetic acid; FLHD, 131-amino acid recombinant full-length heme PAS domain *Bj*FixLH$_{140-270}$; HisTag, histidine tag; LB, Luria broth; TB, terrific broth; MALDI-TOF, matrix-assisted laser desorption ionization time of flight; PAS, domain structure derived from the protein group Per, Arnt, Sim; PCR, polymerase chain reaction; PDB, Protein Data Bank; PMSF, phenylmethanesulfonyl fluoride; AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride; Tris, Trizma Base (Sigma); WSU, Washington State University; NIU, Northern Illinois University; JILA/CU, JILA and University of Colorado.

10.1021/bi701672m CCC: $40.75
Published on Web 01/18/2008

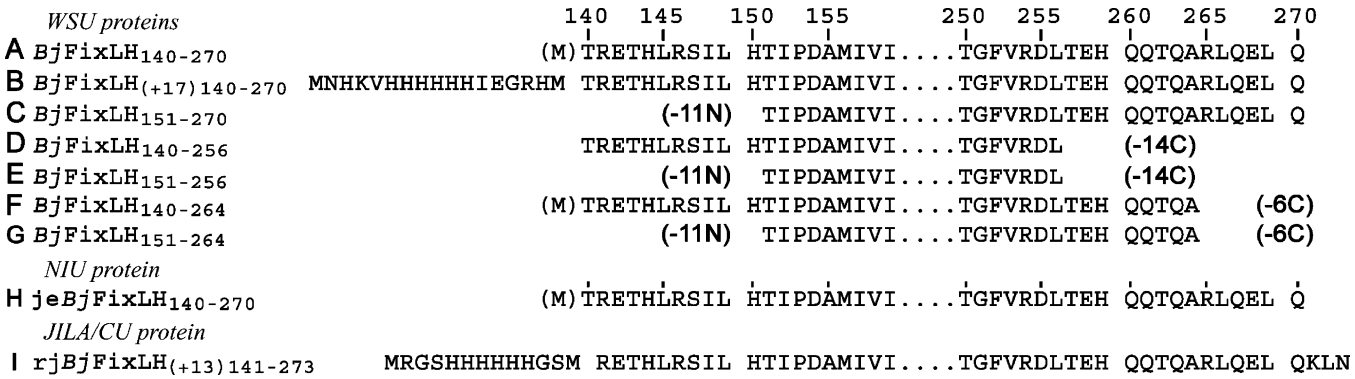

FIGURE 1: Proteins used in this study defined by their specific names and the primary sequences of their N- and C-termini. Terminal sequence truncations are indicated in bold. N-Terminal extensions of proteins B and I result from use of the pCold and pQE30 plasmids. Labeling and ordering are identical to those of Table 1.

transduction can be studied. For *Bj*FixLH, which is the sole subject of this work, the specific recombinant construct used by each of several research groups varies somewhat in primary sequence (*21*, *22*). Our "full-length" heme domain (FLHD), *Bj*FixLH$_{140-270}$ (*21*), is similar in sequence length to the other commonly used *Bj*FixLH parent proteins (*10*−*14*, *16*−*22*).

Initial NMR results for *Bj*FixLH$_{140-270}$ proved to be unexpected, irreproducible, and confusing and were only resolved by mass spectrometry of our recombinant preparations. *Bj*FixLH$_{140-270}$ was found to be unstable over time after isolation, with respect to progressive mass loss (vide infra). We have undertaken an extensive study of *Bj*FixLH$_{140-270}$ and several of its terminal sequence-truncated variants by MALDI-TOF mass spectrometry. Here we report an extensive MALDI-TOF study that documents mass and sequence instability for *Bj*FixLH$_{140-270}$, which may compromise its uncritical use in solution experiments. These detailed results show that the mass instability is not inhibited by a variety of factors, not even by mixed cocktails of protease inhibitors. We have also identified a stable 106-amino acid "heme PAS core", *Bj*FixLH$_{151-256}$, that is resistant to degradation for up to 1 year.

We emphasize that the goal of this study was not to define the mechanism of this degradation. Rather, in view of the apparent extent of the phenomenon, our goal was to report it, to explore its extent and characteristics, and attempt to define its cause. Identifying the mechanism of this phenomenon from degradation pattern analysis will be the topic of a forthcoming study (manuscript in preparation).

Importantly for those working with these proteins, it was also determined that similar recombinant *Bj*FixLH proteins, produced independently in two other, geographically remote laboratories, similarly exhibit mass instability. The results presented here raise the issue of whether this may be a general property of FixLH heme PAS domains longer than 106 amino acids. The reality and extent of this phenomenon are supported by the observation that all published *Bj*FixLH X-ray crystal structures to date show resolved sequences that are significantly shorter than the full sequence of the parent proteins that entered crystallization (vide infra).

EXPERIMENTAL PROCEDURES

BjFixLH Produced at WSU. The gene encoding *Bj*FixL was kindly provided to the Satterlee group (WSU) by H. Hennecke (Eidgenossische Technische Hochschule, Zurich, Switzerland) on plasmid RR28 pRJ7354(FixLJ). The full-length gene was employed as a PCR template for creating the PAS (heme) domain protein, *Bj*FixLH$_{140-270}$, as previously described (*21*).

WSU BjFixLH Variants. The cloned gene for *Bj*FixLH$_{140-270}$ (*21*) was used as a PCR template for creating all of the terminal sequence-truncated variants employed in this work. In addition to the previously described N- and C-terminal primers for *Bj*FixLH$_{140-270}$ (*21*), the following primers were used: (i) *Bj*FixLH$_{151-270}$ (−11N), 5′-AGT CGC ATA TGA CAA TTC CCG AC-3′, which also introduces an NdeI restriction site; (ii) *Bj*FixLH$_{140-256}$ (−14C), 5′-GAG AAG CTT TCA GAG ATC GCG GAC-3′; and (iii) *Bj*FixLH$_{140-264}$ (−6C), 5′-GAG AAG CTT TCA CGC CTG GGT CTG CTG-3′. *Bj*FixLH$_{151-264}$ (−11N/−14C) employed primers i and ii. *Bj*FixLH$_{151-256}$ (−11N/−6C) employed primers i and iii. The required oligonucleotide primers for PCR were synthesized using an Applied Biosystems 380B DNA synthesizer, part of the WSU Laboratory for Biotechnology and Bioanalysis Unit 1 (LBB1).

WSU BjFixLH Production. All of the *Bj*FixLH proteins produced at WSU for this study, summarized in Figure 1A−G and Table 1A−G, were expressed in *Escherichia coli* BL21(DE3λ) or, on occasion, in *E. coli* BL21(DE3λ)pLysS, using isopropyl β-D-thiogalactopyranoside (IPTG) (Fisher Biotech; 1.0 mL of a 1 M solution per liter of growth medium). Two types of expressions were used. (1) Our original procedure involved growth in plasmid pET24a(+) (Novagen) at 37 °C without a histidine tag (*21*) and has continued to be our predominant production method. (2) Production of selected proteins was implemented at lower temperatures using the pCold DNA system (TaKaRa Bio Inc.). This also places a 17-residue N-terminal extension containing a histidine tag on the protein so that the pCold-expressed protein is abbreviated (+17N) in Table 1B and *Bj*FixLH$_{(+17)140-270}$ in Figure 1B. This N-terminal sequence is Met-Asn-His-Lys-Val-His-His-His-His-His-His-Ile-Glu-Gly-Arg-His-Met. The extension is predicted to add 2156.46 Da to the calculated mass of *Bj*FixLH$_{140-270}$. However, as shown in Table 1, this protein has a measured mass that corresponds to a mass increase of only 2025.26 Da. As with most of our other expressed proteins, this mass difference of ~131 Da corresponds to loss of the initial Met in this pCold extension. We have previously shown variable extents of N-terminal Met in our expressed PAS proteins (*21*).

Table 1: *Bj*FixLH Masses Determined by MALDI-TOF Mass Spectrometry

protein	abbreviation	M[a]	M_r measured[b]	SD[c]	95%[d]	n[e]	M_r calcd[f]	Δ[g]	% difference[h]
				WSU					
(A) *Bj*FixLH$_{140-270}$	FLHD[i]	−	14869.13	2.50	0.52	89	14869.84	0.71	0.005
	FLHD	+	15001.08	2.50	0.73	45	15001.04	0.04	0.0003
(B) *Bj*FixLH$_{140-270}$	+17N[j]	−	16894.18	3.83	2.83	7	16895.10	0.92	0.005
	+17N[j]	+	17025.23	2.47	1.21	16	17026.30	1.07	0.006
(C) *Bj*FixLH$_{151-270}$	−11N	−	13525.14	2.12	1.02	19	13525.29	0.15	0.001
	−11N	+	13656.00	1.85	0.76	23	13656.50	0.50	0.004
(D) *Bj*FixLH$_{140-256}$	−14C	−	13177.26	1.96	0.96	16	13178.02	0.76	0.006
	−14C	+	13308.84	1.85	1.00	13	13309.22	0.38	0.003
(E) *Bj*FixLH$_{151-256}$	−11N/−14C	−	11833.92	2.37	1.10	18	11833.47	0.45	0.004
	−11N/−14C	+	11965.69	1.14	0.56	16	11964.67	1.02	0.009
(F) *Bj*FixLH$_{140-264}$	−6C	−	14101.71	1.73	0.80	23	14101.96	0.25	0.002
	−6C	+	NO[k]				14233.16		
(G) *Bj*FixLH$_{151-264}$	−11N/−6C	−	12758.05			1	12757.41	0.64	
	−11N/−6C	+	12887.83			1	12888.60	0.77	
				NIU					
(H) je*Bj*FixLH$_{140-270}$	FLHD	−	14869.33	1.39	2.25	14	14869.84	0.51	0.003
				JILA/CU					
(I) rj*Bj*FixLH$_{141-273}$	+13N[l]	+	16660.00	7.35	2.63	30	16653.87	6.13	0.037

[a] A plus indicates the presence of an N-terminal Met and a minus the absence of an N-terminal Met, as determined by the combination of sequencing and mass analysis. [b] Statistical average mass of the parent neutral molecule determined from MALDI-TOF measurements of the parent +1 ion, corrected for the additional proton. [c] Standard deviation. [d] The 95% confidence interval. [e] Number of measurements in the statistical sample. [f] Mass of the parent neutral molecule calculated from amino acid masses, which were in turn calculated using atomic average isotope masses. [g] Mass difference: |neutral molecule experimental measured mass − calculated mass|. [h] Mass difference expressed as a percentage of the calculated mass. [i] Full-length heme domain. [j] Expressed in the pCold system and carrying an N-terminal histidine tag as described in Experimental Procedures. [k] Not observed in any mass spectrum. [l] Expressed in pQE30 and carrying an N-terminal histidine tag; see Experimental Procedures.

Our earliest methods typically produced mixtures of apoprotein and holoprotein. Those preparations were converted completely to holo-*Bj*FixLH by addition of dissolved protohemin IX (Sigma) at cell lysis. In later preparations, we were able to singularly produce the holoprotein, saturated with heme by supplementing the growth medium with both δ-aminolevulinic acid (Sigma) and ferrous sulfate (Fisher). In these cases, the initial UV−visible spectra, recorded as early as 8−10 min following host cell lysis, showed that *Bj*FixLH proteins were typically isolated predominantly as the oxy−ferrous form. While the initial ratio of oxy to deoxy forms varied from preparation to preparation, the proteins ultimately transformed to the 100% ferrous−oxy form during isolation, consistent with the reported slow oxygen association rates for the structurally related and sequence-related *Sm*FixLH (*23*).

Isolation, Purification, and Storage of WSU BjFixLH. Isolation and purification not involving His-tagged proteins were essentially carried out as previously described (*21*), using a two-column (DEAE and gel filtration) purification. The course of purification was monitored after every major step by SDS−PAGE and UV−visible spectroscopy. Over the course of isolating these proteins for 10 years, small changes in procedure have occurred, including changing from Sephadex G75 gel filtration medium to Sephacryl S100 and S200 media. On occasion, a third gel filtration column was required to complete purification. Protein fractions are now oxidized to the ferriheme state by treatment with potassium ferricyanide (Fisher) prior to gel filtration.

During the ∼48 h period required for purification and isolation, the protein preparations were kept constantly at 4 °C. Histidine-tagged proteins were expressed at 25 °C in the pCold system. Following host cell lysis, they were purified at 4 °C using a Ni^{2+}-NTA (nitrilotriacetic acid) column, prepared and used according to the protocol of TaKaRa Bio Inc. All chromatography columns were washed extensively after each use and then stored in 0.02% sodium azide buffer. This buffer was prewashed away prior to each use. Immediately after the final isolation step, the *Bj*FixLH proteins were sterile filtered through 0.2 μm filters into sterile screw-top tubes for storage at −80 °C. Storage buffer was typically 20 mM Tris (pH 8.0−8.5) and 100 mM NaCl.

Homogeneity and Purity Assessments of WSU Protein. Homogeneity and purity were monitored by SDS−PAGE and UV−visible spectroscopy for all seven recombinant proteins during the course of isolation and purification (vide supra). MALDI-TOF mass spectrometry was carried out on every protein preparation at completion of purification to verify the mass and the mass homogeneity prior to filter sterilization and storage at −80 °C. Frequent amino acid sequencing of ∼12 amino acids at the N-terminus was carried out after the final purification column to confirm the N-terminal sequence identity and to document the variable amounts of the N-terminal methionine.

SDS−PAGE, UV−visible spectroscopy, amino acid sequencing, and protein quantitation were carried out as previously described (*21*). Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry was carried out on a Voyager DE system (Perseptive Biosystems) housed in the WSU Laboratory for Biotechnology and Bioanalysis, Unit 2 (WSU LBB2), as previously described (*21*). Details not previously described are that positive ions were detected in linear mode using delayed extractions (250 ns) with an accelerating voltage of 25000 V. Individual spectra resulted from 32 to 3500 co-added laser shots. Internal references were horse cytochrome *c* and horse myoglobin (both Sigma). An individual protein's mass statistics (Table 1) came from repeated mass determinations on separately made samples, often over many years ($n > 3$; data not shown). Separate MALDI-TOF experiments with one to three internal reference peaks were used in mass calibrations. In all spectra presented here, the parent protein's +1 ion is labeled as $[P + H]^+$.

WSU, NIU, and JILA/CU BjFixLH Aging Experiments. On "day 0" of each aging experiment, freshly isolated or freshly thawed pure solutions of *Bj*FixLH were concentrated or diluted, as required, using a standard stock buffer. This buffer, which was used in every aging experiment, consisted of 20 mM Tris and 100 mM NaCl (pH 8.5). Concentration was carried out using Amicon Ultra-4 centrifugal concentrators. When ligation was desired or the pH had to be changed, the proteins were repeatedly concentrated and diluted with the appropriate buffer. Small samples were taken for both UV−visible spectroscopy and SDS−PAGE. The concentrators were stored in 0.02% sodium azide when not in use and rinsed exhaustively prior to use. After the protein solution had been concentrated, its pH was measured, and then the solution was sterile filtered through either 0.20 μm Nalgene filters or Costar Spin-X centrifuge filters (0.22 μm) directly into sterile tubes and taken immediately into a sterile, laminar flow hood (nuAire, Inc.). All further manipulations were carried out in this hood, including subsequent solution transfers to other sterile tubes and removal of samples for mass spectrometry or spectroscopy. Most room-temperature aging experiments were carried out in this hood. Three aging experiments were carried out in the oxygen-free environment of an MBraun Unilab inert atmosphere glovebox. Typical aging experiments extended to between 30 and 105 days and consisted of 10−70 μL samples that were allowed to age in capped sterile Eppendorf-type tubes that were themselves contained in sterile screw-top tubes. Sterile water (100 μL) was placed into each screw-top tube prior to each sealing to minimize evaporation of the aging solutions. Samples (1 μL) were removed from individual sealed tubes for time sequence MALDI-TOF mass spectrometry evaluation of aging progress.

Strict attention was paid to sterility and sterile technique throughout. For example, the water used in making all buffers was from a Barnstead Epure three-stage filtering system, fitted with a 0.20 μm sterile hollow fiber fourth-stage filter. Sterile transfers out of the laminar flow hood to the Unilab glovebox were carried out by placing individual samples in capped Eppendorf-type tubes into larger sterile screw-top tubes, capping them, and then taking them directly to the glovebox entry port where the outside of the tube was washed with 90% ethanol prior to executing the evacuation/entry cycles and entry into the glovebox. Aerobic and anaerobic comparison aging experiments could not be carried out with the samples continuously exposed to their respective atmospheres because of the problem of solvent evaporation over the time scale of the aging experiments. To compensate, both types of aging were carried out in closed tubes, but these tubes of protein solutions were initially equilibrated with their respective atmospheres for ~2 h before initial capping. Then, each tube was uncapped for approximately 30 min each day during the aging experiment to equilibrate the sample with the ambient atmosphere.

Protease Inhibitors for Isolation, Purification, and Aging of WSU BjFixLH. All protease inhibitors were purchased from Sigma and used according to the manufacturer's instructions. For isolations, the protease inhibitors were added immediately after *E. coli* cell lysis at initiation of the isolation. Generally, these isolations were carried out in the presence of only irreversible protease inhibitor AEBSF, 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride, or (in early work) PMSF, phenylmethanesulfonyl fluoride. However, the effects of various mixtures of protease inhibitors were also tested on several isolations. Protease mixtures, including commercial protease cocktails, were repeatedly used in aging experiments. In addition to AEBSF (or PMSF), the additional inhibitors used were antipain, aprotinen, bestatin, benzamidine-HCl, chymostatin, ethylenediaminetetraacetic acid (EDTA), ϵ-amino-*n*-caproic acid, E64, *N*-ethylmaleimide, *trans*-epoxysuccinyl-L-leucylamido(4-guanidino)butane, *N*-(α-rhaminopyranosyloxyhydroxyphosphinyl)-Leu-Trp, leupeptin, and pepstatin A. Two commercial (Sigma) protease inhibitor cocktails were also used, either individually or in combination in aging experiments: p8465 (Sigma), for all bacterial extracts, and p2714-1BT (Sigma), for general use. When protease inhibitors were to be included in aging experiments, they were added to the concentrated protein solutions either before or after the final concentration and before sterile filtering (see above).

BjFixLH Produced at NIU. The heme domain of BjFixL was cloned by amplifying codons 140−270 of the *Bj*FixL gene using genomic DNA of *B. japonicum* as the template using standard PCR technology. This protein has a sequence identical to that of the WSU *Bj*FixLH$_{140-270}$, and it is denoted as je*Bj*FixLH$_{140-270}$ in Figure 1H and Table 1H. Primers were designed to introduce NdeI and EcoRI restriction sites at the 5′ and 3′ ends of the amplified fragment, respectively (for details, see ref *60*). The amplified gene fragment was ligated into pET-24b(+) (Novagen). This recombinant plasmid, called pET-24b(+)/BjFixL$_H$, was used to express *Bj*FixLH$_{140-270}$ in *E. coli*.

The pET-24b(+)/BjFixL$_H$ plasmid was transformed into *E. coli* BL21(DE3), and a 1 L culture was grown at 37 °C in TB medium containing kanamycin at 20 μg/mL. Expression was induced with 1 mM IPTG at an OD$_{600}$ of 1−1.2. The induced culture was grown overnight at 28 °C. Cells were harvested and lysed in a French press. The supernatant containing *Bj*FixLH$_{140-270}$ was applied to a Sephadex G-75 gel filtration column and eluted with a buffer containing 20 mM potassium phosphate and 100 mM NaCl (pH 7.5). Fractions containing the protein were pooled and applied to a DEAE-Sepharose (FastFlow) column and eluted with a 50 to 500 mM NaCl gradient over 300 mL. The purity of the protein was determined by SDS−PAGE and UV−vis spectroscopy.

BjFixLH Produced at JILA/CU. The *Bj*FixL gene was provided by H.-M. Fischer (Eidgenossische Technische Hochschule) on pRJ7349, which was used as the template for PCR amplification of the heme domain, *Bj*FixLH$_{141-270}$. The forward and reverse primers included BamHI and HindIII restriction sites, respectively: forward, 5′-TAGGATCCATGCGCGAGACCCACC-3′; reverse, 5′-TATAAGCTTTTGCAGTTCCTGGAGAC-3′. Both the purified PCR product and a pQE30 (Qiagen) plasmid were doubly digested with these enzymes and gel purified, and the *Bj*FixLH segment was then ligated to pQE30.

Protein expression was carried out as follows. *E. coli* BL21(DE3) cells were transformed with the BjFixLH/pQE30 plasmid. A single-colony transformant (from the LB/ampicillin plate) was grown overnight, and 45−50 mL of the culture grown from the single colony was used to inoculate 2 L (in a 6 L flask) of medium containing 1 mg/mL ampicillin, supplemented with 80 mg/L δ-aminolevulinic acid and 10 μM $FeSO_4$. This mixture was grown at 37 °C with

vigorous shaking until the OD_{600} reached 0.5−0.6, at which time IPTG was added (to 1 mM), the temperature reduced to 30 °C, and the shaker speed reduced to 180 rpm. Cells were harvested after ~3 days and pellets kept frozen until they were needed.

Because the pQE30 plasmid confers a six-histidine tag at the N-terminus of *Bj*FixLH, purification was performed in a single-affinity chromatography step. Resuspension and cell lysis (by lysozyme and Triton X) were performed in 20 mM sodium phosphate and 500 mM NaCl (pH 7.8). The cell lysate (after centrifugation and 0.45 μm filtration) was poured onto a Ni-NTA (Qiagen) column, rinsed with the same buffer, and then washed with buffers of decreasing pH, down to pH 5.3. The red/brown *Bj*FixLH fraction was eluted with the buffer at pH 4.0; the *Bj*FixLH-containing eluant was dialyzed against 20 mM Tris buffer and 150 mM NaCl (pH 8.5). Aliquots were frozen until needed. The translated amino acid sequence of this protein [called rj*Bj*FixLH$_{(+13)141-273}$ in Figure 1I and abbreviated (+13) in Table 1I] is MRGSHHHHHHGSMRETHLRSILHTIPDAMIVIDGHGIIQLFSTAAERLFGWSELEAIGQNVNILMPEPDRSRHDSYISRYRTTSDPHIIGIGRIVTGKRRDGTTFPMHLSIGEMQSGGEPYFTGFVRDLTEHQQTQARLQELQKLN.

Bioinformatics. (1) Secondary Structure Analyses. Structure files from the Protein Data Bank (*24*) were visualized with DS ViewerPro (Accelrys), as shown in Figures 9 and 10. DS ViewerPro and visual inspection are the least quantitative secondary structure assessments shown in Figure 9. Secondary structure analysis of the structure data (Figure 9a) was implemented using DSSP (*25*) on the heme domain structure 1DRM (*16*). Secondary structure prediction software was applied for analyzing and predicting secondary structure components directly from the primary sequence of *Bj*FixLH$_{140-280}$ (*18*; Figure 9d−g). Four programs were used: PredictProtein programs (i) Profile Network Prediction Heidelberg, PROF_sec, (ii) a subset of PROF_sec, which has a claimed predictive accuracy greater than 82%, SUB_sec (*26*, *27*), (iii) the Advanced Protein Structure Prediction Server, APSSP2, a neural network-based software (*28*), and (iv) PSIPRED (*29*, *30*).

(2) Disorder Structure Prediction. Seven programs were used to predict disordered regions directly from the primary sequence of *Bj*FixLH$_{140-280}$. These were PONDR (*31*, *32*); disEMBL (*33*), which contains three different predictors [(i) loops/coils definition, (ii) hot loops definition, and (iii) PDB file "remark-465" definition]; DISOPRED (*34*); FOLDINDEX (*35*); IUPRED (*36*, *37*); RONN (*38*); and GLOBPLOT (*39*). The results are presented in Figure 9h−n.

RESULTS AND DISCUSSION

Description of Proteins Used. To precisely specify the sequence identity of the several variant *Bj*FixLH proteins that were employed in this work, we continue to employ the nomenclature proposed previously (*21*). The nine recombinant *Bj*FixLH proteins used in this work are identified by this nomenclature, and they appear in identical order in Table 1A−I and Figure 1A−I. Experiments described in the following used exclusively *Bj*FixLH in the Fe^{3+} (oxidized, ferric) state because partial oxidation occurs during the course of purification [similar to *Sm*FixLH (*23*)], otherwise resulting in mixed oxidation states of the protein that complicates chromatography.

Figure 1 defines the varying N-terminal and C-terminal primary sequences of the nine proteins used. These variations in termini are their differentiating features and represent the focus of this work. Mass data for all of these proteins are listed in Table 1 and confirm the identity of each protein by virtue of the excellent agreements between calculated and experimental masses. The experimental masses given in Table 1 are those of the apoprotein, since heme is lost during the course of making MALDI-TOF samples.

Figure 1A shows primary sequence details for the termini of our 131-amino acid parent full-length heme domain, *Bj*FixLH$_{140-270}$. We previously published mass spectrometry and SDS−PAGE assessments of freshly isolated and purified preparations of this protein (*21*), and new experimental mass data are given in Table 1A. This sequence also represents a separate protein preparation made at NIU, denoted je*Bj*FixLH$_{140-270}$, which is entry H in Table 1 and Figure 1.

Figure 1B represents the full-length heme domain expressed in the pCold plasmid, which produces a 17-amino acid N-terminal extension. As described in Experimental Procedures, our recombinant proteins exhibit variable extents of N-terminal methionine derived from the ATG translation initiator. *E. coli* contains methionine aminopeptidase, which acts to remove N-terminal Met (*40*), but its efficiency with overexpressed proteins is often less than 100%. In our experience with recombinant PAS heme domains, this typically leads to two parent mass peaks differing by *m*/*z* 131, indicating a purified mixture consisting of proteins with and without N-terminal Met. In Table 1, these are indicated with plus and minus signs, respectively, in the column labeled M (Met). Other sequences are explicitly defined in Figure 1 and Table 1.

Figure 1I (Table 1I) represents the full-length heme domain produced at JILA/CU, which contains a 13-amino acid N-terminal extension as a result of expression in the pQE30 plasmid. It is also four amino acids longer at the C-terminus than *Bj*FixLH$_{140-270}$.

MALDI-TOF Mass Spectrometry of Expressed BjFixLH Proteins. Each of the nine *Bj*FixLH proteins shown in Figure 1 that were made for this work was subjected to mass analysis by MALDI-TOF mass spectrometry for the purposes of characterization. Mass determinations were used to confirm correct protein production and to assess the homogeneity and purity of every preparation of recombinant proteins made at WSU during the past 9 years. These data, presented in Table 1, came from preparations of unaged proteins that were freshly isolated and purified or from frozen stock solutions that had been frozen immediately after the final purification step (which we call day 0 spectra).

The mass agreement in all cases but one is equal to or better than 0.009% (by mass), which provides confirmatory evidence of the appropriate sequence identity of each recombinant protein. Each day 0 spectrum of the WSU proteins (Figure 1A−G and Table 1A−G) showed peaks in only the *m*/*z* 2500−20000 region because of the parent protein of interest (and any added internal references). Spectra of these proteins over that larger *m*/*z* region are not shown, but spectra over *m*/*z* regions of the +1 ions are presented in Figures 2−7. The larger range mass spectrum of *Bj*FixLH$_{140-270}$ has previously been published (*21*), and all other "fresh" preparations gave similar spectra, all of which indicated a high degree of sample homogeneity.

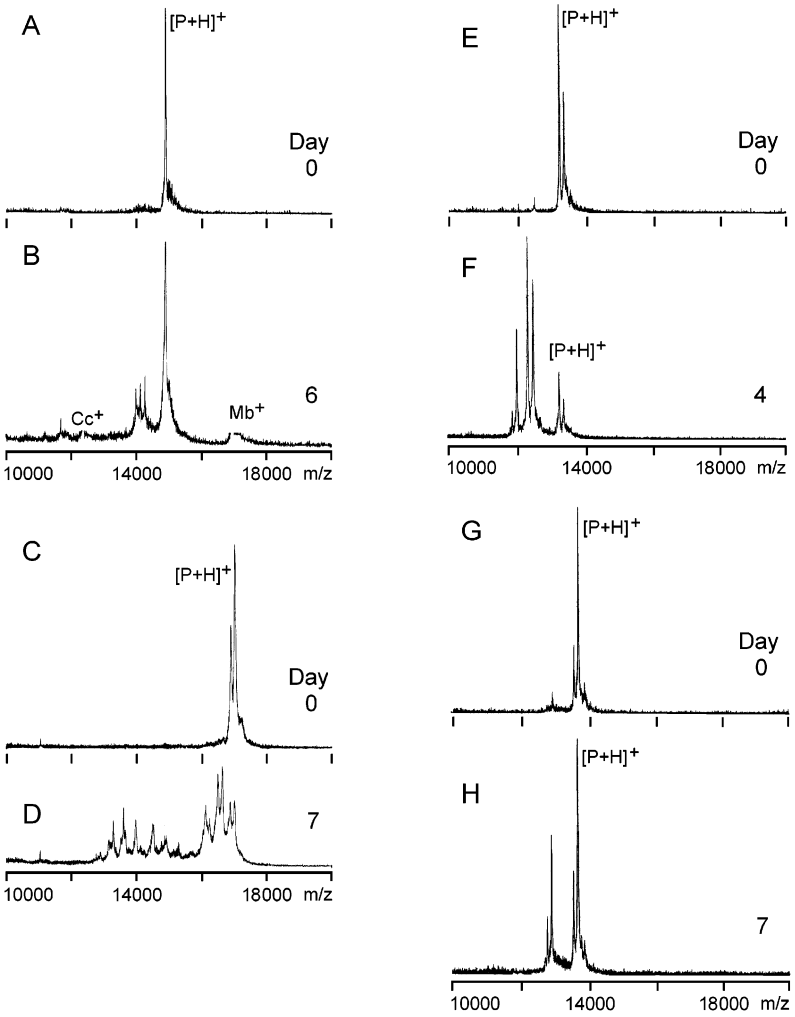

FIGURE 2: MALDI-TOF mass spectra of recombinant WSU *Bj*FixLH proteins made for this work. Pairs of mass spectra are shown for each of the freshly prepared individual proteins (day 0) and those proteins after subsequent days of aging (day 4, 6, or 7) as indicated on each spectrum: (A and B) full-length recombinant heme domain, *Bj*FixLH$_{140-270}$; (C and D) *Bj*FixLH$_{140-270}$ expressed in the pCold system that adds a 17-amino acid HisTag extension to the N-terminus; (E and F) *Bj*FixLH$_{140-256}$ (−14C); and (G and H) *Bj*FixLH$_{151-270}$ (−11N). Two parent peaks in day 0 spectra that are separated by 131 Da represent populations with or without N-terminal methionine.

Mass Spectrometry Assessment of Aging Experiments for WSU Recombinant BjFixLH Proteins. Data such as that presented in Figures 2−4 show that *Bj*FixLH preparations from three (geographically distinct) separate laboratories each display time-dependent room-temperature mass loss. We have found that this mass loss can be detected upon close inspection of even freshly isolated and purified samples after just a few hours, and as the data show, it is substantial after 4−7 days.

Figure 2A shows the MALDI-TOF +1 ion mass spectrum of freshly isolated (i.e., day 0) recombinant *Bj*FixLH$_{140-270}$, where $[P + H]^+$ indicates the parent ion. As previously demonstrated (*21*), the spectrum in the +1 ion region displays a single peak at *m*/*z* 14869, corresponding to the calculated mass for the *Bj*FixLH$_{140-270}$ sequence without a terminal Met (Table 1). (In this preparation, the fraction of the preparation carrying the N-terminal Met is negligible.) Figure 2B shows that after 6 days of room-temperature aging under sterile conditions several lower-mass peaks appear around *m*/*z* 14000, which must be derived from degradation of the parent protein.

A similar pair of mass spectra are shown in panels C and D of Figure 2 for the N-extended *Bj*FixLH$_{(+17)140-270}$ variant (Figure 1B and Table 1B). The day 0 spectrum of the freshly thawed, freshly isolated protein shown in Figure 2C displays two parent peaks indicating that a fraction of the expressed protein retains the N-terminal Met. The N-terminal extension results in a parent (+1) ion mass (Figure 2C and Table 1B) higher than that for the protein in Figure 2A. Figure 2D shows the results of sterile aging at room temperature for 7 days. This protein is clearly very susceptible to extensive age-related mass loss, as witnessed by the many peaks at *m*/*z* values lower than that of the parent protein at day 7. Several peaks occur as low as *m*/*z* ∼13000−14000, representing mass losses of up to ∼3500 Da. In our experience, this type of aging is typical and will be described in detailed analyses of *Bj*FixLH mass degradation patterns (manuscript in preparation).

Also shown in Figure 2 are similar pairs of day 0 and aged spectra for the recombinant −14C variant, *Bj*FixLH$_{140-256}$ (Figure 2E,F), and the −11N variant, *Bj*FixLH$_{151-270}$ (Figure 2G,H). These two variants represent the stable C-terminus and, separately, the stable N-terminus, respectively, of the 106-amino acid stable PAS core protein (see further); hence, neither would be expected to lose further mass from their truncated terminus. Nevertheless, both of these proteins do

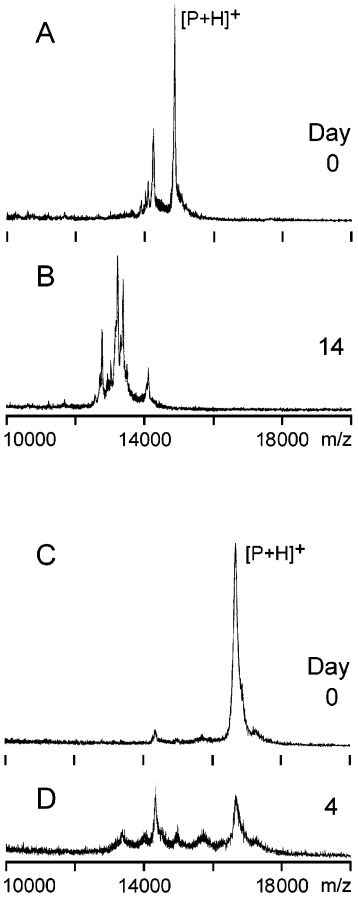

FIGURE 3: MALDI-TOF mass spectra of recombinant *Bj*FixLH proteins made in other laboratories and used in aging experiments. *Bj*FixLH$_{140-270}$ was made at NIU: (A) freshly prepared, as received sample at day 0, and (B) after aging for 14 days. *Bj*FixLH$_{141-273}$ was made at JILA/CU. This protein was expressed with a 13-amino acid extension to the N-terminus and a three-amino acid extension to the C-terminus: (C) freshly prepared, as received sample at day 0, and (D) after aging for 4 days.

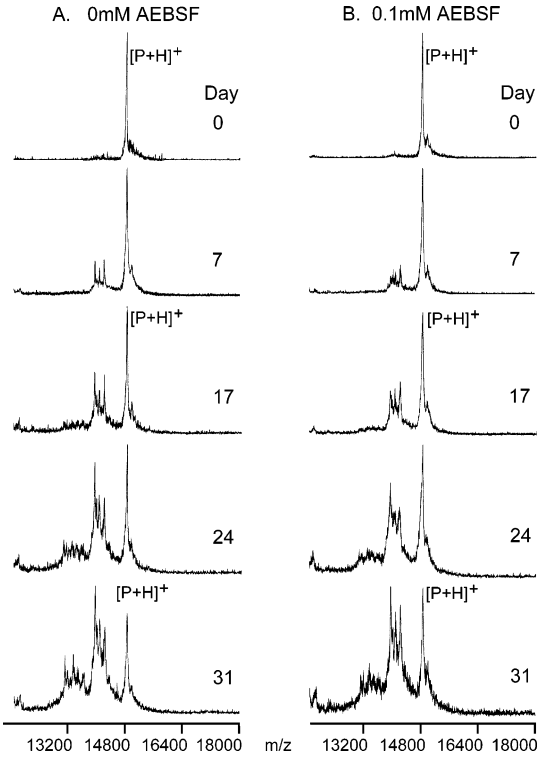

FIGURE 4: MALDI-TOF mass spectra of 1 mM *Bj*FixLH$_{140-270}$ aged in a sterile hood at room temperature without (A) and with (B) the protease inhibitor AEBSF. (A) No AEBSF added and aging for up to 31 days, as indicated on the spectra. (B) Aging in the presence of a 10 mol % inhibitor concentration (0.1 mM AEBSF).

lose further mass upon aging as shown by the multiple low *m*/*z* peaks. This behavior indicates that both termini can be degraded. This is confirmed by combination sequencing and aging experiments that will be subsequently published.

Mass Spectrometry Assessment of Preparations and Aging Experiments for Recombinant BjFixLH Proteins from NIU and JILA/CU. To determine if the observed degradation behaviors were unique to *Bj*FixLH preparations from the WSU laboratory, we performed identical aging experiments with *Bj*FixLH preparations obtained from NIU (Figure 1H and Table 1H) and JILA/CU (Figure 1I and Table 1I). The results of MALDI-TOF measurements on je*Bj*FixLH$_{140-270}$ expressed, purified, and isolated in the Erman laboratory at NIU are shown in panels A and B of Figure 3. Figure 3A shows the MALDI-TOF spectrum of the "as received"(at WSU) protein (day 0). This spectrum showed two peaks in the +1 ion range, one at full mass [*m*/*z* 14869 (Table 1)] that is labeled as the parent peak and one major peak at lower mass, revealing that degradation had already occurred. After 14 days of aging (Figure 3B) there was no parent peak left, only lower mass peaks because of degraded proteins derived from the parent protein.

A second protein was independently obtained from the Jimenez laboratory (JILA/CU), rj*Bj*FixLH$_{(+13)141-273}$ and the similar pair of mass spectra that were recorded are shown in panels C and D of Figure 3. Figure 3C shows the as received protein right after thawing, with the major peak demonstrating the mass of the parent protein (*m*/*z* 16660). The spectrum also exhibited minor lower-mass peaks in the +1 ion region. Significant low-mass peaks developed as early as 4 days later (Figure 3D).

Possible Causes of Mass Loss. We have carried out a total of 21 aging experiments, under diverse conditions, with as many as 10 different protein preparations in each experiment simultaneously, and in all cases, proteins longer than the 106-amino acid core PAS heme domain (*Bj*FixLH$_{151-256}$) always displayed mass loss. The obvious possible causes of this phenomenon are (i) bacterial contamination, (ii) copurification of a protease, or (iii) an inherent property of heme PAS domains. Significant effort over 7 years was devoted in attempting to identify the cause. We consider each.

(1) Bacterial Contamination. Bacterial contamination is ruled out for these reasons. Preparations and aging experiments carried out since 2001 included sterile filtering of protein preparations immediately following the final purification step and prior to freezing (−80 °C) for storage. Sample solutions were again sterile filtered after being thawed for use and prior to initiation of aging in sealed sterile tubes. Some aging experiments were carried out in the presence of added NaN_3 and KCN, both inhibitors of bacterial growth. Aging experiments that did not require an inert atmosphere were carried out in a sterile laminar flow hood. Visual inspection throughout aging experiments never revealed any solid in the sample, or solution cloudiness, as would be expected in the case of bacterial contamination. At the end of aging experiments, all samples were also subjected to

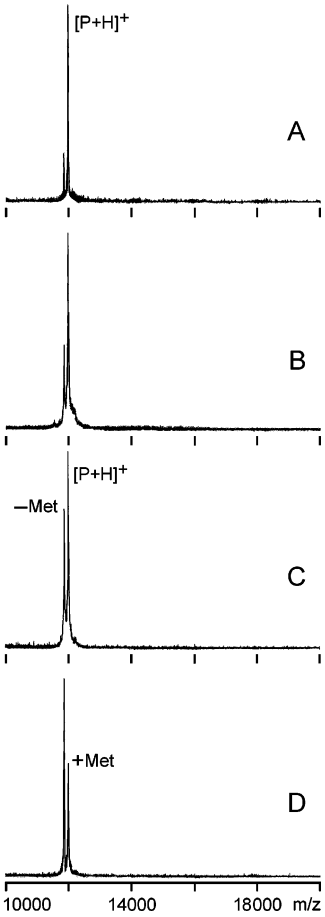

FIGURE 5: MALDI-TOF mass spectra of the highly truncated, 106-amino acid, recombinant *Bj*FixLH$_{151-256}$ (core PAS domain) after varying extents of aging. Two parent peaks are observed in the preparation (*21*), indicating that the mixture contained expressed proteins with (+) and without (−) Met at the N-terminus prior to T_{140}. (A) Freshly prepared sample at day 0 and (B) sample at day 32 of sterile aging. (C) Separate freshly prepared sample at day 0, before NMR experiments, and (D) after 1 year in the NMR tube, primarily kept at 4 °C, but used for NMR experiments at room temperature during that year.

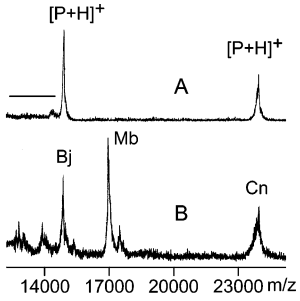

FIGURE 6: MALDI-TOF spectra showing aging of *Bj*FixLH$_{140-270}$ in the presence of casein. (A) Protein mixture at day 0 with the parent +1 ion peaks noted. The horizontal line defines the *m*/*z* region where lower-mass peaks appeared upon degradation. (B) Day 4 of aging. Bj = *Bj*FixLH$_{140-280}$, Cn = casein, and Mb = myoglobin internal reference for MALDI-TOF.

olfactory and visual assessment, and neither indicated contamination.

(2) Protease Contamination. A persistent protease could not be ruled out as being a cause, although several experiments suggest that it is unlikely. The steps we took to rule out protease contamination follow. (i) Since 1999, all

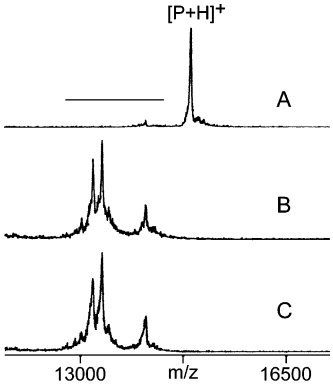

FIGURE 7: MALDI-TOF spectra illustrating the independence of mass degradation of *Bj*FixLH$_{140-280}$ with respect to oxygen concentration. (A) Spectrum at day 0 of aging showing the parent +1 ion peak and very little degradation. The horizontal line indicates the lower-mass region in which peaks of degraded proteins appeared. (B) Day 16 of aging in the absence of O_2. (C) Day 16 of aging in the presence of atmospheric O_2.

preparations were treated with at least protease inhibitor PMSF or AEBSF at isolation. (ii) The *E. coli* strain employed in our production [BL21(DE3)] was engineered to be deficient in Lon and OmpT proteases (*41*, *42*). *E. coli* Lon and ClpAP proteases are known to recognize protein degrons in vivo according to the "N-end Rule" (*43*−*45*). However, all of our *Bj*FixLH proteins have sequences beginning with either Met or Thr, each of which confers stability against this type of degradation (*43*). Further, C-terminal sequences of all of our proteins also confer stability against *E. coli* tail specific protease, Tsp, in vivo and in vitro (*46*).

(iii) We have carried out repeated aging experiments in the presence of commercial protease inhibitor cocktails, single protease inhibitors, or combinations of protease inhibitors, none of which has stopped the observed mass loss. An example of our results is shown in Figure 4, which compares sterile aging for up to 31 days for 1 mM solutions of *Bj*FixLH$_{140-270}$ in the absence (Figure 4A) and presence (Figure 4B) of covalent protease inhibitor AEBSF. The figures indicate that the pattern of mass spectra is essentially independent of the presence of AEBSF when its concentration is as high as 10% (0.1 mM) of that of the protein. Similar aging spectra are found when an AEBSF concentration of 1% of that of the protein (0.01 mM AEBSF) was used (not shown).

(iv) A 106-amino acid variant, *Bj*FixLH$_{151-256}$ (−11N/−14C), truncated at each terminus in comparison to *Bj*FixLH$_{140-270}$ is highly resistant to mass loss. This is illustrated in the MALDI-TOF spectra shown in Figure 5. This protein is another example of variable N-terminal labeling with Met, as indicated in Figure 5. Figure 5A shows a mass spectrum of *Bj*FixLH$_{151-256}$ at day 0 in one aging experiment, whereas Figure 5B shows the same preparation after sterile, room-temperature aging for 32 days. No low-mass peaks appear upon aging, indicating no mass degradation occurs for *Bj*FixLH$_{151-256}$ (−11N/−14C). Figure 5C is the mass spectrum of a separate preparation of *Bj*FixLH$_{151-256}$ taken at a different day 0 (3 years prior to the previous experiment), before its use for ∼12 days in room-temperature NMR experiments. Figure 5D is a mass spectrum taken 1 year later. During that year, the sample had been primarily stored in a refrigerator (4 °C) but had spent several weeks

in the NMR magnet at room temperature. Clearly, this truncated core PAS heme domain is highly stable.

(v) Six experiments were carried out in which *Bj*FixLH$_{140-270}$ was co-aged in solutions together with various combinations of several other proteins. The proteins used were commercial horse heart cytochrome *c*, commercial horse heart myoglobin, commercial casein, native recombinant *Saccharomyces cerevisiae* cytochrome *c* peroxidase, and recombinant (H52L) cytochrome *c* peroxidase (both of the latter prepared using the same *E. coli* strain employed here for *Bj*FixLH). In all cases, the other proteins showed no degradation or mass loss over the time in which substantial degradation was observed for *Bj*FixLH$_{140-270}$. Casein, a highly disordered protein, is known to be particularly susceptible to proteases. The results of co-aging *Bj*FixLH$_{140-270}$ with casein are shown in Figure 6. Figure 6A shows a MALDI-TOF spectrum of the protein mixture at initiation of the aging experiment (day 0), whereas Figure 6B shows the spectrum (with Mb added as an internal mass reference) of the mixture after it had been sterilely aged for 4 days. In Figure 6B, low-mass peaks indicating mass loss in *Bj*FixLH$_{140-270}$ appear in the *m*/*z* 12500−14500 region (indicated with a horizontal straight line), but no peaks that would indicate casein degradation occur between *m*/*z* ∼15000 and ∼23000. Thus, mass loss for *Bj*FixLH$_{140-270}$ is observed under conditions when it is not observed for casein or for other co-aged proteins (not shown). (vi) Three experiments were conducted in which mixtures of the four proteins described in part (v) were aged separately but in parallel identically with *Bj*FixLH$_{140-270}$. The results, again, were that *Bj*FixLH$_{140-270}$ degraded while the other proteins did not. Further, fractional precipitation during or after purification did not alter the degradation behavior.

None of these experiments can absolutely rule out protease contamination as the cause of the observed spontaneous protein degradation. If it is a protease, it is persistent and resists even mixtures of up to two complete protease inhibitor cocktails. It is also unusual in that it must also be present in preparations from two geographically remote, independent laboratories, one employing a substantially different expression system.

Chemistry of Heme PAS Domains? If the reproducible spontaneous mass degradation of *Bj*FixLH proteins longer than 106 amino acids is due to chemistry associated with the protein itself, it can be differentiated for analysis into (a) heme-based chemistry and (b) chemistry not related to heme. Two types of experiments were carried out to assess heme's role in mass degradation, aging experiments of *Bj*FixLH$_{140-270}$.

Aging experiments were initiated with apo-*Bj*FixLH$_{140-270}$, but it was so unstable in solution at pH 8.5 that only low-quality spectra were obtained for samples aged for more than 7 days (Supporting Information). These spectra do show the presence of low-mass peaks similar to those presented above, and similar to those of simultaneously aged heme-reconstituted holoprotein. Although preliminary, such results suggest that the apoprotein and holoprotein undergo similar mass loss. Further effort is underway to improve the apo aging experiment.

The second type of experiment consisted of *Bj*FixLH$_{140-270}$ aging in the presence of excess amounts of selected heme ligands appropriate for a ferriheme protein: CN^-, N_3^-, and imidazole. Imidazole and CN^- were previously demonstrated to be ligands for *Bj*FixLH (*47*, *48*). It is well-known that in heme proteins the locus of heme-based chemistry is the heme iron ion, which functions as a ligand-binding site, as a redox center, and in acid−base chemistry. Thus, it was hypothesized that ligating the heme iron ion would alter heme redox or acid−base chemistry that could be the source of the observed polypeptide degradation. None of these ligands inhibited the observed mass loss (data not shown), which supports the idea from the apo-*Bj*FixLH$_{140-270}$ aging experiment that heme is not involved in the aging mass loss.

A related test is whether the protein mass loss is affected by oxygen, which in the presence of heme iron ion could undergo redox chemistry to produce potent polypeptide disruptors. This was tested using an inert atmosphere glovebox to provide an oxygen-free environment for aging ferric *Bj*FixLH$_{140-270}$, while oxygen exposure came simply from the atmosphere. Two pairs of aging experiments were simultaneously performed, in which one sample of the pair was aged in the inert atmosphere box while the other sample in the pair was aged in the sterile hood. Representative results are shown in Figure 7. Figure 7A is the mass spectrum of *Bj*FixLH$_{140-270}$ at day 0. Figure 7B is the spectrum at day 16 of aging in the absence of oxygen. Figure 7C is the spectrum at day 16 of aging in the presence of oxygen. From the similarity of the two sets of spectra, there is no obvious oxygen dependence on aging. Essentially identical degradation in the two samples is reflected by the essentially identical pattern of low-*m*/*z* peaks that develop with time (Figure 7). Taken together, these three types of aging experiments (apoprotein, heme ligation, and oxygen independence) suggest that there is no heme dependence upon aging.

This leaves polypeptide-based chemistry as a possible source for the aging mass loss. Among many possibilities, the presence of a domain with protease activity is perhaps the most obvious one, but the co-aging experiments (vide supra) showed that proteins in solution with *Bj*FixLH$_{140-270}$ were not degraded; hence, such a domain in *Bj*FixLH$_{140-270}$ would have to be extremely self-specific. It would also have to be part of the "stable core" (*Bj*FixLH$_{151-256}$) since the termini are lost in this process. We have not yet been able to identify a "protease-like" domain or structural motif in *Bj*FixLH. Another possibility may be that the termini of *Bj*FixLH$_{140-270}$ are simply chemically unstable under our solution conditions. Solution chemical studies in progress aim to further elucidate these aspects of *Bj*FixLH$_{140-270}$ behavior.

Having established the mass instability of three independent preparations of the longer *Bj*FixLH heme domains, we find a significant possibility exists that this phenomenon is general. Published X-ray crystal structures are relevant to this point.

Sequences of Crystal Structures. Fifteen X-ray crystal structures of *Bj*FixL heme domains in various heme oxidation and ligation states have been published to date (*16*, *18*, *47*−*50*). Figure 8 depicts the N-terminal and C-terminal sequences for each of these crystal structure proteins in the same manner as in Figure 1. These structures are grouped according to the sequence of the parent protein that was committed to crystallization.

Nine of the structures (*16*, *47*−*49;* Figure 8B−F, PDB entries 1DRM−1DP8) derive from a 130-amino acid parent

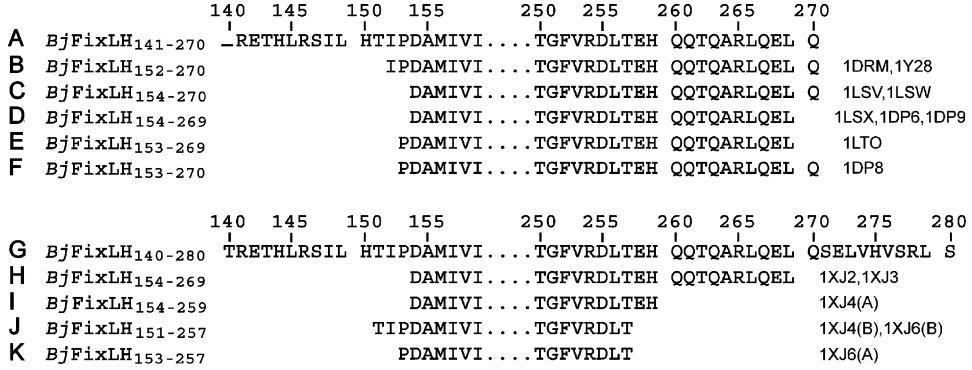

FIGURE 8: N- and C-terminal primary sequences of X-ray crystal structure proteins. (A) Full sequence of the parent protein used for crystallization that produced structures with the resolved sequences shown in rows B−F (*16*, *47*−*49*). (G) Full sequence of the parent protein used for crystallization that produced structures with the resolved sequences shown in rows H−K (*18*, *50*). PDB entries are given to the right of each structure sequence.

FIGURE 9: Secondary structure analysis (a−g) and disorder prediction (h−n) of *Bj*FixLH$_{140-280}$ (*18*) referenced to its primary sequence. The top two rows are the numbering system and explicit primary sequence, respectively. Secondary structure analyses and predictions follow. (a) Analysis of crystal structure 1DRM (*16*) using DSSP (*25*): E represents (extended) β-strand, H α-helix, T hydrogen-bonded turn, and S bend. (b and c) Graphical interpretations of secondary structures of (b) *Bj*FixLH$_{152-270}$ [1DRM (*16*)] and (c) *Bj*FixLH$_{154-269}$ [1XJ2 (*18*)] by DS ViewerPro (Accelrys). (d and e) Analysis using the PredictProtein (*26*, *27*) programs (PROF_sec = Profile Network Prediction Heidelberg, and SUB_sec = a subset of PROF_sec), where E represents β-strand, L loop, and H helix. (f) APSSP2 [Advance Protein Structure Prediction Server (*28*)] in which H represents helix, E strand, and C (random) coil. (g) PSIPRED in which H, E, and C have the meanings as in part f (*29*, *30*). Blank spaces indicate no prediction made. Protein disorder predictions follow. D indicates disorder, and blank spaces indicate ordered regions. (h) Prediction by PONDR VL-XT (*31*, *32*). (i) Prediction by disEMBL, with three predictors: −1 = loops/coils, −2 = hot loops, and −3 = remark-465 (*33*). (j) DISOPRED (*34*). (k) FoldIndex (*35*). (l) IUPRED (*36*, *37*). (m) RONN (*38*). (n) GLOBPLOT (*39*).

protein, *Bj*FixLH$_{141-270}$, shown in Figure 8A. The resolved sequences of both termini that are reported in the X-ray structures for each of the nine structures are given in Figure 8B−F, along with the corresponding PDB entry for each. The nine structures (PDB entries 1DRM−1DP8) are each missing substantial portions of their N-termini (−11N to −13N) compared to the parent protein that entered crystallization (Figure 8A). In addition, two of the structures, panels D and E, show a single missing C-terminal amino acid.

Six structures (*18*, *50*; Figure 8H−K, PDB entries 1XJ2−1XJ6) derive from a substantially longer parent protein sequence consisting of 141 amino acids, *Bj*FixLH$_{140-280}$,

shown in Figure 8G. Terminal sequences of those six structures are shown in Figure 8H−K. Compared to the 141-amino acid parent, all of these structures (PDB entries 1XJ2−1XJ6) are also missing substantial portions of their N-terminal sequences (−11N to −14N). In addition, these structures are also missing 11, 21, or 23 C-terminal amino acids. The hallmark of all 15 of these structures is that they contain a resolved core sequence in the range of 105−120 amino acids in length, similar to our stable 106-amino acid *Bj*FixLH core protein (vide supra). All X-ray structures have substantial unresolved or missing terminal segments compared to the proteins that entered crystallization. There is no actual structural characterization of their termini, which means that the extant X-ray structures of *Bj*FixL heme domains represent proteins that are significantly smaller than the parent proteins from which they derive, approximately 10−25% smaller.

Some of the PDB files explicitly acknowledge this situation by listing missing residues, indicating undetected electron density. However, in most cases, the PDB file comments and the publications reporting these structures do not mention the missing amino acids and focus on only the resolved core structures. So the following question arises: Are the missing residues in these X-ray structures physically absent, consistent with the mass spectrometry data presented here, or are the termini of these proteins structurally intact, but disordered in the crystal? Absent electron density could be attributed to either of these situations (*51*). Within the protein disorder literature, missing X-ray coordinates are associated with disordered regions (*31*−*33*, *38*, *51*−*54*), as well as with missing residues. Disordered regions are also associated with hypersensitivity to protease digestion (*32*, *51*), which could produce missing residues.

With respect to this, we note that the protein manifestation of "disorder" seems not to have a definition upon which there is universal agreement. Dunker and co-workers define disorder as lack of tertiary (three-dimensional) structure (*31*, *32*), while others define it to mean "...lack of regular secondary structure and a high degree of flexibility in the polypeptide chain" (*33*, *55*). Another definition is that disordered regions are "...random coil-like, molten globule-like, or somewhere in between" (*51*). Ferren et al. define disorder as "...the lack of stable secondary and tertiary structure..." (*54*).

In view of the foregoing, we subjected *Bj*FixLH$_{140-280}$ and *Bj*FixLH$_{140-270}$ to a suite of secondary structure predictions and disorder predictions (*25*−*39*) in an attempt to obtain insight into the extent of secondary and tertiary structure in the N- and C-terminal regions that are missing from X-ray structures. Such an assessment might help resolve the source of the missing residues and provide structural information about the *Bj*FixLH termini.

Secondary Structure. A set of secondary structure prediction programs were applied to the primary sequence of *Bj*FixLH$_{140-280}$ and yielded the results shown in Figure 9a−g. The top two rows in Figure 9 present the *Bj*FixLH$_{140-280}$ primary sequence and numbering system, as in Figure 8, except that in this case the entire sequence is explicitly shown.

In Figure 9, entry a represents the DSSP (*25*) secondary structural analysis of PDB entry 1DRM (Figure 8), in which E represents a β-strand, H represents helix, and T and S

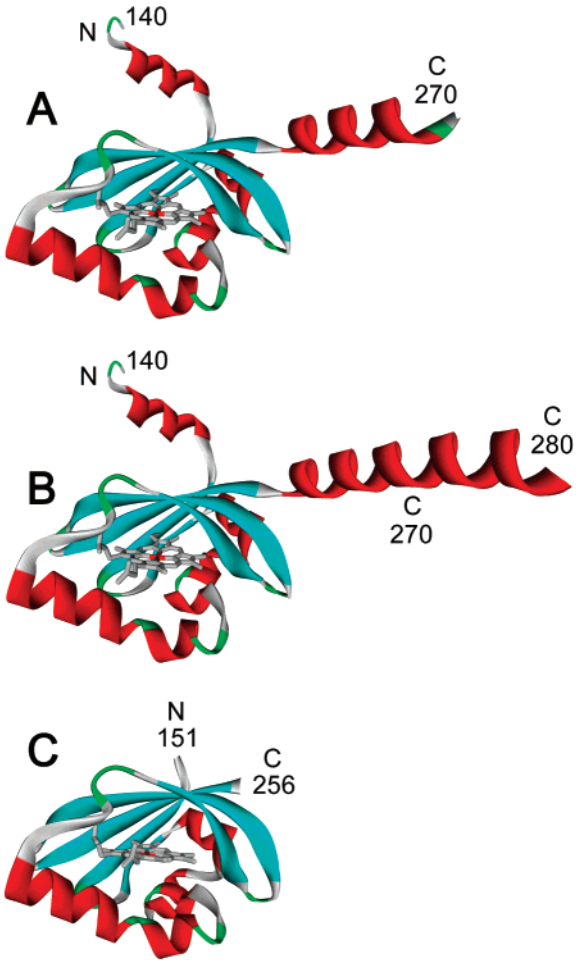

FIGURE 10: (A) Schematic structure representation of *Bj*FixLH$_{140-270}$ (Figure 1A and Table 1A) built from the structure denoted as 1DRM (*16*; Figure 8B). It includes N-terminal residues 140−151 that are missing from the structure, which are shown as predominantly a helix corresponding to the results shown in Figure 9. (B) Schematic structure representation of *Bj*FixLH$_{140-280}$ (*18*), the parent protein for the group of structures described in Figure 8H−K (*18*), built from 1DRM (*16*; Figure 8B). It includes N-terminal residues 140−151 and C-terminal residues 269−280 that are missing from all extant heme domain structures. (C) Model of the 106-amino acid, stable, truncated, core PAS heme domain, *Bj*FixLH$_{151-256}$ (Figure 1E and Table 1E).

represent turns and bends, respectively. DSSP analyzes actual structures, interpreting them in terms of its criteria for defining secondary structure elements. It is not a prediction program. Rows b and c are the secondary structure elements of actual *Bj*FixLH structures defined by the visualization program DS ViewerPro (Accelrys) for PDB entries 1DRM (*Bj*FixLH$_{141-270}$) and 1XJ2 (*Bj*FixLH$_{140-280}$), respectively (Figure 8). E and H have the same definitions as in DSSP (row a), but L (loop) represents turns and bends. Rows a−c are in excellent agreement in defining the secondary structure motifs that are present and combine to form a structure-based consensus.

Rows d−g consist of secondary structure predictions from types of software that use protein primary sequence (not structure) as their predictive basis. Row d and e report the results of two PredictProtein (*26*, *27*) programs. In these programs, E, H, and L have the meanings given above. Row f is another secondary structure prediction using the Advanced Protein Structure Prediction Server (*28*), in which E and H again refer to β-strand and helix, respectively, and C

indicates a predicted (random) coil region. Row g displays PSIPRED (*29*, *30*) data in which E, H, and C are defined in a manner identical to that of APSSP2 (row f).

Over the PAS core primary sequence region, $A_{155}-L_{269}$, the structure-based secondary structure definitions (Figure 9 a−c) are virtually identical to the primary sequence-based secondary structure predictions (Figure 9d−g). This agreement suggests that the sequence-based predictive programs are reliable predictors of *Bj*FixLH$_{140-280}$ secondary structure, which in turn, gives confidence in these programs' assignments of helices to the termini, for which no structure has been determined (Figure 9d−g). The consensus view of these four programs was that the N-terminus ($T_{140}-D_{154}$) is a helix and that the C-terminus beyond H_{259} is also a helix. These results lead to the conclusion that the termini maintain well-defined secondary structure.

Disorder Prediction. *Bj*FixLH$_{140-280}$ and *Bj*FixLH$_{140-270}$ were also subjected to analysis by seven different sequence-based disorder prediction programs, consisting of nine different criteria. The results are shown in Figure 9h−n. In contrast to the high level of agreement in secondary structure predictions, there is nearly no agreement among the disorder predictions. Considering the termini, only PONDR VL-XT (*31*, *32*; Figure 9h) and IUPRED (*36*, *37*; Figure 9l) predicted disorder at the N-terminus (indicated with a D). The other seven predictors did not assign disorder there, indicating it to be ordered.

Four of the nine predictors indicated substantial disorder at the C-terminus. However, only two, disEMBL-3 (remark-465 definition; *33*) and RONN (*38*), predicted disorder to the extent necessary to account for the missing C-terminal segments in PDB entries 1XJ2 and 1XJ3 (Figure 8). Only RONN (*38*) predicted disorder to the extent required to account for PDB entries 1XJ4 and 1XJ6 (Figure 8).

Structural Models. Figure 10 presents model structures built from the original PDB structure file, 1DRM (*16*), incorporating the results of the consensus secondary structure prediction from Figure 9d−g. These are only hypothetical model structures designed to transmit a structural sense of the parent protein structures (Figure 1D−F). To aid this, the secondary structure elements are displayed by color. Details, such as terminal helix orientation, are not intended to be structurally accurate, but rather suggestive. Figure 10A depicts a model of our full-length heme domain, *Bj*FixLH$_{140-270}$ (Figure 1A), which is also essentially the parent protein (Figure 8A) for the structures shown in Figure 8B−F. Figure 10B represents a model of *Bj*FixLH$_{140-280}$ (Figure 8G), the parent protein for the structures represented by the sequences in Figure 8H−K. Figure 10C represents the highly truncated, highly stable, core PAS heme domain *Bj*FixLH$_{151-256}$ (Figure 1E), the MALDI-TOF spectra of which are presented in Figure 5.

SUMMARY AND CONCLUSIONS

The experimental results presented here show that several recombinant *Bj*FixL heme domains, whose primary sequences are significantly longer than 106 amino acids, are unstable (as isolated) with respect to mass loss upon standing at room temperature. Only a 106-amino acid stable heme PAS core, *B*jFixLH$_{151-256}$, is stable to aging mass loss. It was demonstrated that this degradation is observed reproducibly for three separate types of heme domain preparations from geographically remote labs, one of which used a production protocol significantly different from the other two. The spontaneous mass loss of recombinant *Bj*FixLH$_{140-270}$ begins immediately after purification, and it was not inhibited by protease inhibitors, the presence or absence of O_2, or alteration of the heme ligation state of the holoprotein. Preliminary data showed mass loss in aging the apoprotein as well. We have found that while sample refrigeration reduces the rate of mass degradation it does not eliminate it. Freezing at −80 °C has stopped the degradation in samples kept for 9−10 years.

Our experimental data make protease contamination unlikely as the source of the aging mass loss, attributing it rather to inherent polypeptide instability. Such instability is likely to have manifested itself in the *Bj*FixLH X-ray crystal structures. The consensus secondary structure prediction of helices at the *Bj*FixLH termini taken together with the inconsistent predictions of disorder in the termini suggests to us that disorder as the cause of the missing X-ray structure segments is unlikely. It seems more likely to us that the missing segments are physically absent. It seems likely that PAS heme domain mass degradation to a stable, minimal core structure is general.

In related studies (manuscripts in preparation), we have detected comparable mass losses in two other heme PAS domains: (i) FixLH from *S. meliloti* (*Sm*FixLH; formerly *R. meliloti*) and (ii) the *E. coli* direct oxygen sensor heme domain, *Ec*DosH. In contrast, preliminary results indicate that the globin-like heme-sensing domain derived from the *Bacillus subtilis* aerotaxis protein, HemAT, is stable to mass loss for at least 1 month. We emphasize that the data presented here do not rule out protease activity as the cause of this phenomenon.

Observation of non-protease mass loss in heme-containing proteins is not without precedent, as we have previously documented a specific C-terminal dipeptide truncation in *Neisseria meningitidis* heme oxygenase (*56*, *57*). That dipeptide was similarly missing from the *Nm*HemO crystal structure (*58*, *59*).

The results presented here should stimulate use of high-resolution mass determinations in assessment of all extant *Bj*FixLH preparations. We volunteer our help in that effort. Proof of mass integrity of *Bj*FixLH proteins before and after experimental use would seem to us to be required henceforth. In this respect, we used MALDI-TOF mass spectrometry to assess *Bj*FixLH$_{140-270}$ and *Bj*FixLH$_{151-256}$ samples that had been used in photothermal experiments (*15*) and found them to be fully intact following the experiments. If mass degradation is found to be a general problem, then experimental handling protocols will need to be adopted to minimize degradation during experiments, as we did in our photothermal work. Controlled experiments will also be needed to determine whether there are any functional consequences.

ACKNOWLEDGMENT

We thank Professor Dr. Hauke Hennecke and Professor Dr. Hans-Martin Fischer, both at the Eidgenossische Technische Hochschule, for providing the *Bj*FixL gene to J.D.S. and R.J., respectively. R.J. is a staff member in the Quantum

Physics Division of NIST. J.D.S. acknowledges the expertise of Harold Goldberg, M.D., without whom this work would not have been achieved. We gratefully acknowledge use of the RCSB PDB, which is supported by funds from the National Science Foundation (NSF), the National Institute of General Medical Sciences (NIGMS), the Office of Science, Department of Energy (DOE), the National Library of Medicine (NLM), the National Cancer Institute (NCI), the National Center for Research Resources (NCRR), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the National Institute of Neurological Disorders and Stroke (NINDS), and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). We gratefully acknowledge use of the WSU LBB2 mass spectrometer facility and the peptide synthesis capability of WSU LBB1.

SUPPORTING INFORMATION AVAILABLE

Sequential MALDI-TOF spectra of a simultaneous aging experiment of apo-*Bj*FixLH$_{140-270}$ and heme-reconstituted *Bj*FixLH$_{140-270}$ (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Virts, E. L., Stanfield, S. W., Helinski, D. R., and Ditta, G. S. (1988) Common regulatory elements control symbiotic and microaerobic induction of nifA in *Rhizobium meliloti*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 3062−3065.
2. David, M., Daveran, M.-L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., and Kahn, D. (1988) Cascade regulation of nif gene expression in *Rhizobium meliloti*, *Cell 54*, 671−683.
3. Reyrat, J. M., David, M., Blonski, C., Boistard, P., and Batut, J. (1993) Oxygen-regulated *in vitro* transcription of *Rhizobium meliloti* nifA and fixK genes, *J. Bacteriol. 175*, 6867−6872.
4. de Philip, P., Batut, J., and Boistard, P. (1990) *Rhizobium meliloti* FixL is an oxygen sensor and regulates *R. meliloti nifA* and *fixK* genes differently in *Escherichia coli*, *J. Bacteriol. 172*, 4255−4262.
5. Sciotti, M.-A., Chanfon, A., Hennecke, H., and Fischer, H.-M. (2003) Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*, *J. Bacteriol. 185*, 5639−5642.
6. Mesa, S., Ucurum, Z., Hennecke, H., and Fischer, H. M. (2005) Transcription activation *in vitro* by the *Bradyrhizobium japonicum* regulatory protein FixK2, *J. Bacteriol. 187*, 3329−3338.
7. Monson, K., Ditta, G. S., and Helinski, D. R. (1996) The oxygen sensor protein, FixL, of *Rhizobium meliloti*, *J. Biol. Chem. 270*, 5243−5250.
8. Anthamatten, D., and Hennecke, H. (1991) The regulatory status of the fixL- and fixJ-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*, *Mol. Gen. Genet. 225*, 38−48.
9. Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. (1991) A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*, *Nature 350*, 170−172.
10. Monson, E. K., Weinstein, M., Ditta, G. S., and Helinski, D. R. (1992) The FixL protein of *Rhizobium meliloti* can be separated into a heme-binding oxygen-sensing domain, and a functional C-terminal kinase domain, *Proc. Natl. Acad. Sci. U.S.A. 89*, 4280−4284.
11. Gilles-Gonzalez, M. A., Gonzalez, G., Perutz, M. F., Kiger, L., Marden, M. C., and Poyart, C. (1994) Heme-based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation, *Biochemistry 33*, 8067−8073.
12. Gilles-Gonzalez, M. A., Gonzalez, G., and Perutz, M. F. (1995) Kinase activity of the oxygen sensor FixL depends on the spin state of its heme iron, *Biochemistry 34*, 232−236.
13. Gilles-Gonzalez, M. A., and Gonzalez, G. (2005) Heme-based sensors: Defining characteristics, recent developments, and regulatory hypotheses, *J. Inorg. Biochem. 99*, 1−22.
14. Gilles-Gonzalez, M. A., and Gonzalez, G. (2004) Signal transduction by heme-containing PAS-domain proteins, *J. Appl. Physiol. 96*, 774−783.
15. Miksovska, J., Suquet, C., Satterlee, J. D., and Larsen, R. W. (2005) Characterization of conformational changes coupled to ligand photodissociation from the heme binding domain of FixL, *Biochemistry 44*, 10028−10036.
16. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998) Structure of a biological oxygen sensor: A new mechanism for heme-driven signal transduction, *Proc. Natl. Acad. Sci. U.S.A. 95*, 15177−15182.
17. Lukat-Rodgers, G. S., Rexine, J. L., and Rodgers, K. R. (1998) Heme speciation in alkaline ferric FixL and possible tyrosine involvement in the signal transduction pathway for regulation of nitrogen fixation, *Biochemistry 37*, 13543−13552.
18. Key, J., and Moffat, K. (2005) Crystal structures of deoxy and CO-bound BjFixLH reveal details of ligand recognition and signaling, *Biochemistry 44*, 4627−4635.
19. Miyatake, H., Mukai, M., Park, S.-Y., Adachi, S.-I., Tamura, L., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T., and Shiro, Y. (2000) Sensory mechanism of oxygen sensor FixL from *Rhizobium meliloti*: Crystallographic, mutagenesis and resonance Raman spectroscopic studies, *J. Mol. Biol. 301*, 415−431.
20. Liebl, U., Bouzhir-Sima, L., Negrerie, M., Martin, J.-L., and Vos, M. H. (2002) Ultrafast ligand rebinding in the heme domain of the oxygen sensors FixL and Dos: General regulatory implications for heme-based sensors, *Proc. Natl. Acad. Sci. U.S.A. 99*, 12771−12776.
21. Suquet, C., Savenkova, M., and Satterlee, J. D. (2005) Recombinant PAS-heme domains of oxygen sensing proteins: High level production and physical characterization, *Protein Expression Purif. 42*, 182−193.
22. Satterlee, J. D., Suquet, C. M., Savenkova, M. I., and Lian, C. (2003) Proton NMR characterization of recombinant ferric heme domains of the oxygen sensors FixL and Dos: Evidence for protein heterogeneity, *ACS Symp. Ser. 858*, 244−257.
23. Gonzalez, G., Gilles-Gonzalez, M. A., Rybak-Akimova, E. V., Buchalova, M., and Busch, D. H. (1998) Mechanisms of autoxidation of the oxygen sensor FixL and *Aplysia* myoglobin: Implications for oxygen-binding heme proteins, *Biochemistry 27*, 10188−10194.
24. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank, *Nucleic Acids Res. 28*, 235−242.
25. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2637.
26. Rost, B., Yachdav, G., and Liu, J. (2004) The PredictProtein Server, *Nucleic Acids Res. 32*, W321−W326.
27. Liu, J., and Rost, B. (2004) Sequence-based prediction of protein domains, *Nucleic Acids Res. 32*, 3522−3530.
28. Raghava, G. P. S. (2002) APSSP2: A combination method for protein secondary structure prediction based on neural network and example based learning, *CASP5 A-132* (http://www.imtech.res-.in/raghava/apssp2/).
29. Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London, *Nucleic Acids Res. 33*, W36−W38.
30. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices, *J. Mol. Biol. 292*, 195−202.
31. Li, X., Romero, P., Rani, M., Dunker, A. K., and Obradovic, Z. (1999) Predicting protein disorder for N-, C- and internal regions, *Genome Inf. 10*, 30−40.
32. Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001) Sequence complexity of disordered protein, *Proteins: Struct., Funct., Genet. 42*, 38−48.
33. Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003) Protein disorder prediction: Implications for structural proteomics, *Structure 11*, 1453−1459.
34. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life, *J. Mol. Biol. 337*, 635−645.
35. Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I., and Sussman, J. L. (2005) FoldIndex©: A simple tool to predict whether a given protein sequence is intrinsically unfolded, *Bioinformatics 21*, 3435−3438.

36. Dosztanyi, Z., Csizmok, P. T., and Simon, I. (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins, *J. Mol. Biol. 347*, 827−839.
37. Dosztanyi, Z., Csizmok, P. T., and Simon, I., (2005) IUPRED: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics 21*, 3433−3434.
38. Yang, Z. R., Thomson, R., McNeil, P., and Esnouf, R. M. (2005) RONN: The bio-basis function neural network technique applied to the detection of natively disordered regions in proteins, *Bioinformatics 21*, 3369−3376.
39. Linding, R., Russell, R. B., Neduva, V., and Gibson, T. J. (2003) GlobPlot: Exploring protein sequences for globularity and disorder, *Nucleic Acids Res. 31*, 3701−3708.
40. Liao, Y.-D., Jeng, J.-C., Wang, C.-F., Wang, S.-C., and Cheng, S.-T. (2004) Removal of N-terminal methionine from recombinant proteins by engineered *E. coli* methionine aminopeptidase, *Protein Sci. 13*, 1802−1810.
41. Studier, W. F., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol. 185*, 80−119.
42. http://www.stratagene.com/homepage/default.aspx, search "BL21".
43. Varshavsky, A. (1992) The N-end rule, *Cell 69*, 725−735.
44. Graciet, E., Hu, R.-G., Piatkov, K., Rhee, J. H., Schwarz, E. M., and Varshavsky, A. (2006) Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen, *Proc. Natl. Acad. Sci. U.S.A. 103*, 3078−3083.
45. Wang, K. H., Sauer, R. T., and Baker, T. A. (2007) ClpS modulates, but is not essential for bacterial N-end rule degradation, *Genes Dev. 21*, 403−408.
46. Keiler, K. C., and Sauer, R. T. (1996) Sequence determinants of C-terminal substrate recognition by the Tsp protease, *J. Biol. Chem. 271*, 2589−2593.
47. Gong, W., Hao, B., and Chan, M. K. (2000) New mechanistic insights from structural studies of the oxygen-sensing domain of *Bradyrhizobium japonicum* FixL, *Biochemistry 39*, 3955−3062.
48. Hao, B., Isaza, C., Arndt, J., Soltis, M., and Chan, M. K. (2002) Structure-based mechanism of O_2 sensing and ligand discrimination by the FixL heme domain of *Bradyrhizobium japonicum*, *Biochemistry 41*, 12952−12958.
49. Dunham, C. M., Dioum, E. M., Tuckerman, J. R., Gonzalez, G., Scott, W. G., and Gilles-Gonzalez, M. A. (2003) A distal arginine in oxygen-sensing heme-PAS domains is essential to ligand binding, signal transduction, and structure, *Biochemistry 42*, 7701−7708.
50. Key, J., Srajer, V., Pahl, R., and Moffat, K. (2007) Time-resolved crystallographic studies of the heme domain of the oxygen sensor FixL: Structural dynamics of ligand rebinding and their relation to signal transduction, *Biochemistry 46*, 4706−4715.
51. Rani, M., Romero, P., Obradovic, A., and Dunker, A. K. (1998) Annotation of PDB with respect to "disordered regions" in proteins, *Genome Inf. 9*, 240−241.
52. Dunker, A. K., Oldfield, C. J., Meng, J., Romero, P., Yang, J. Y., Obradovic, Z., and Uversky, V. N. (2007) Intrinsically disordered proteins: An update, *Proc. 7th IEEE Int. Conf. Bioinf. Bioeng. 1*, 49−58.
53. Radivojac, P., Iakoucheva, L. M., Oldfield, C. J., Obradovic, Z., Uversky, V. N., and Dunker, A. K. (2007) Intrinsic disorder and functional proteomics, *Biophys. J. 92*, 1439−1456.
54. Ferron, F., Longhi, S., Canard, B., and Karlin, D. (2006) A practical overview of protein disorder prediction methods, *Proteins: Struct., Funct., Bioinf. 65*, 1−14.
55. Wright, P., and Dyson, H. J. (1999) Intrinsically unstructured proteins: Re-asessing the protein structure-function paradigm, *J. Mol. Biol. 293*, 321−331.
56. Liu, Y., Ma, L.-H., Satterlee, J. D., Zhang, X., Yoshida, T., and La Mar, G. N. (2006) Characterization of the spontaneous "aging" of the heme oxygenase from the pathological bacterium *Neisseria meningitidis* via cleavage of the C-terminus in contact with the substrate: Implications for functional studies and the crystal structure, *Biochemistry 45*, 3875−3886.
57. Liu, Y., Ma, L.-H., Zhang, X., Yoshida, T., Satterlee, J. D., and La Mar, G. N. (2006) ^{1}H NMR study of the influence of hemin vinyl → methyl substitution on the interaction between the C-terminus and substrate and the "aging" of the heme oxygenase from *Neisseria meningitidis*: Induction of active site structural heterogeneity by a two-fold symmetric hemin, *Biochemistry 45*, 13875−13888.
58. Schuller, D. J., Zhu, W., Stojiljkovic, I., Wilks, A., and Poulos, T. L. (2001) Crystal structure of heme oxygenase from the gram negative pathogen *Neisseria meningitidis* and a comparison with mammalian heme oxygenase-1, *Biochemistry 40*, 11552−11558.
59. Friedman, J. M., Lad, L., Deshmukh, R., Li, H. Y., Wilks, A., and Poulos, T. L. (2003) Crystal structures of the NO- and CO bound heme oxygenase from *Neisseria meningitidis*: Implications for O_2 activation, *J. Biol. Chem. 278*, 34654−34659.
60. Bidwai, A. K. (2003) Ph.D. Dissertation, Northern Illinois University, DeKalb, IL.

BI701672M