Journal of Inorganic Biochemistry 105 (2011) 609–615

Contents lists available at ScienceDirect

Journal of Inorganic Biochemistry

journal homepage: www.elsevier.com/locate/jinorgbio

Origins of aging mass loss in recombinant N-terminus and C-terminus deletion mutants of the heme-PAS biosensor domain *Bj*FixLH$_{140–270}$

James D. Satterlee *

Department of Chemistry, Washington State University, Pullman, WA 99164–4630, USA

ARTICLE INFO

Article history:
Received 26 August 2010
Received in revised form 20 January 2011
Accepted 21 January 2011
Available online 3 February 2011

Keywords:
MALDI-TOF
FixLH
*Bj*FixLH
Bradyrhizobium japonicum
Heme domain
Oxygen sensor

ABSTRACT

Nine recombinant FixL heme domains from *Bradyrhizobium japonicum* previously were shown to exhibit mass instability independent of many environmental factors (J.D. Satterlee, C. Suquet, A. Bidwai, J. Erman, L. Schwall, R. Jimenez, Biochemistry 47 (2008) 1540–1553). Two of those recombinant proteins were produced in remote laboratories. Mass losses begin appearing at completion of isolation and comprise a substantial proportion of samples within 1–3 days of storage and handling. Thus, degradation occurs during the time frame of experiments and crystallization. Detailed understanding of this instability is desired in order to formulate stable heme-PAS sensor domains for experimentation and for a mechanistic interpretation. However, mass spectra of the full length heme-PAS domain, *Bj*FixLH$_{140–270}$, are complex by 1–3 days following isolation due to broad features and a high density of overlapping peaks, so that individual peak assignments are at present ambiguous. This stymies direct, quantitative interpretation of the source of the observed mass losses. To solve this dilemma amino-terminal primary sequencing and MALDI-TOF (Matrix Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry monitoring of three terminal variants of *Bj*FixLH$_{140–270}$ have been achieved. The working hypothesis, that the experimentally observed mass losses originate in the PAS protein sequence termini, has been substantiated. This establishes a basis for interpreting the more complex results from aging full length *Bj*FixLH$_{140–270}$.

1. Introduction

Nine separate recombinant constructs of *Bj*FixLH$_{140–270}$, representing the recombinant heme-PAS sensing domain of the larger two-domain protein *Bj*FixL, were recently shown to be unstable with respect to mass loss beginning immediately after purification [1]. Since heme-PAS domains as well as non-heme PAS domains are now widely studied, characterization of this mass loss has been continued in order to fully understand its source. In that previous work it was found that the degradation occurred nearly identically in individual recombinant *Bj*FixLH$_{140–270}$ proteins made independently in three geographically remote laboratories. The published work [1] also revealed that mass loss occurred for apo-*Bj*FixLH$_{140–270}$ and for two *Bj*FixLH proteins with N-terminal extended sequences.

*Bj*FixLH$_{140–270}$, our recombinant full length heme domain (FLHD), is the O_2-sensing domain of *Bj*FixL, which is found in the symbiotic bacterium *Bradyrhizobium japonicum* (*Bj*). FixL itself is one of a pair of proteins that comprise FixLJ, which is a two component regulator of bacterial physiological responses to the micro-oxic environment within the root nodule [2–5], where the symbiote is localized. *Sm*FixLJ is the analogous protein in the related symbiote *Sinorhizobium meliloti* [2,6–8]. In both species, FixL consists of ~500 amino acids, divided into a heme sensing domain and an enzymatic kinase domain [9,10].

FixLs from both species function similarly, in which dissociation of heme-ligated O_2 at low oxygen partial pressure triggers an interdomain signal that activates the kinase domain (both *in vivo* and *in vitro*), which in concert with ATP, phosphorylates, thus activating, the response regulator-transcription factor FixJ [11–13]. FixJ initiates a regulatory cascade involving several genes needed for low-oxygen adaptation and nitrogen fixation [2,5,14]. With respect to the degree of direct regulation of nitrogenase gene activation, the two species differ [3–5], and in this work we concentrate specifically on *Bj*FixLH, the Bradyrhizobium japonicum sensing domain.

Interest in FixL derives from its biosensor role in which it was observed that ferrous heme ligation status in the FixL sensing (heme) domain (FixLH) was communicated to the FixL kinase domain, indicating an interdomain signal [15–20]. One can imagine a process in which O_2 dissociating from ferrous heme initiates a signal within the heme domain (intradomain signal) that evolves into the kinase domain as an interdomain signal, resulting in kinase-domain activation. Just such a heme domain signal has been proposed [21] based on Photoacoustic Calorimetry results following photodissociation of CO-*Bj*FixLH$_{140–270}$ [20]. Ideas about FixL signaling have developed out of extensive previous structural and dynamics studies [9,11–28, and references therein] using the two-domain recombinant *Sm*FixL and

* Fax: +1 509 335 8867.
E-mail address: hemeteam@wsu.edu.

0162-0134/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jinorgbio.2011.01.015

*Bj*FixL (~50 and ~40 kD, respectively) as well as the smaller recombinant heme domains, *Sm*FixLH and *Bj*FixLH (each ~15 kD).

All structural and most chemical and biochemical studies to-date have been performed on the recombinant heme domains because they are smaller, highly soluble and well characterized. [11,22–27]. In contrast, no three dimensional structure of a two-domain FixL protein has yet been reported.

Many different recombinant *Bj*FixLH and *Sm*FixLH heme domains have been constructed for use in different laboratories [e.g. 1,28,29 and references therein]. Irreproducible ^{1}H NMR spectra for several presumed pure preparations of recombinant *Bj*FixLH$_{140–270}$ lead this laboratory to subsequent mass analyses, which in turn revealed that multiple mass peaks had developed over time in the NMR solutions. For NMR experiments stable forms of *Bj*FixLH are required, which has lead to further mass spectrometry and aging studies in attempts to define a stable *Bj*FixLH [1,this work]. The prior [1] working hypothesis, that the observed mass losses originate in both the N-terminus and the C-terminus is addressed here using controlled aging of selected deletion mutants that were monitored by primary sequencing and mass determinations by MALDI-TOF mass spectrometry.

The results provide an interpretive basis for understanding the much greater mass spectral complexity resulting from aging degradation of the full length heme domain, *Bj*FixLH$_{140–270,}$ and its his-tagged variants. It is likely that this same paradigm applies to similar aging mass losses that have been observed by this laboratory for other PAS heme sensing domains, including *Sinorhizobium meliloti* FixL (*Sm*FixLH) and *E. coli* Direct Oxygen Sensor (*Ec*DosH) (unpublished data).

2. Experimental

2.1. General methods

The production techniques, including various primers, isolation and purification procedures and all chemicals and procedures for the recombinant native and variant *Bj*FixLH proteins used in this work have been described in detail previously [1,28]. Primary sequencing of the amino termini was carried out as described previously [28]. Initially purified preparations of the expressed parent protein of the three variants described herein, *Bj*FixLH$_{140–270}$, are mass homogeneous as indicated by MALDI-TOF (Matrix Assisted Laser Desorption Ionization Time of Flight) mass spectrometry [1,28]. All solutions were made using 18 MΩ water (Barnstead) that was also filter sterilized. Purity and homogeneity assessments of each protein preparation by mass spectrometry were carried out as previously described [1,28]. Pertinent to the question of the origin of the observed mass losses in all the recombinant *Bj*FixL heme domains from this laboratory is the fact that the isolation and purification procedures always included addition of protease inhibitor cocktail (Sigma P2714; P8465) and separately AEBSF (Sigma A8456) (early work was with PMSF) when the *E. coli* BL21 cells were lysed at the start of isolation. The protein solutions were refreshed with protease inhibitors after the last (2nd) purification column. Purified solutions were filter sterilized, refreshed with protease inhibitors and stored at −80 °C. Details of MALDI-TOF mass spectrometry carried out on a Voyager DE system (Perseptive Biosystems) in linear mode were identical to those previously described [1,28], including the matrix (sinapic acid, Sigma). Mass calibrations utilized 3-point (equine cytochrome $c+1$ ion and equine myoglobin $+1$ and $+2$ ion) and 2-point (myoglobin only) internal references. Sufficient signal-to-noise required 32–1396 shots per spectrum. Two detection mass ranges were used: 2500 m/z–20,000 m/z and 7000 m/z–18,000 m/z, leading to digital resolutions of 1.54–3.47 m/z per data point. Aging experiments were carried out as previously described [1] in which concentrated samples (20–100 μL total volume) to be aged, taken from freezer stocks, were refreshed with protease inhibitors then filter sterilized into sterile 1.5 mL capped conical tubes for aging. These were immediately transferred to a sterile hood for aging in larger capped sterile plastic tubes. Stock aging solutions ranged between 300 and 800 μM in heme domain. One to five μL samples from individual aging tubes were periodically extracted inside the sterile hood for mass spectrometry and, in two instances, for amino-terminal primary sequencing during the course of aging experiments.

2.2. Mass matching to identify low-mass peaks

Uniquely identifying the sequence of degrading *Bj*FixLH proteins from peaks in mass spectra was achieved by initially assuming that in the case of the −14C variant, mass losses detected by mass spectrometry resulted from amino acid losses only at the N-terminus (intact terminus). For the −14C variant this assumption was guided by amino terminal sequencing (Table 1). It was also validated experimentally when unique mass matches for all low-mass peaks in the aging MALDI-TOF spectra for all three variants were found (*vide infra*). Unique peak assignments to specific protein amino-acid sequences required: (1) accurate experimental determination of mass and (2) accurate calculation of masses for all possible sequences, then (3) comparing the two data sets. For our purposes sufficiently accurate mass determinations were possible with the Voyager DE spectrometer (See Tables 2–4). The digital resolution of the experiments presented here ranged between 3.47 and 1.54 m/z per data point. Consequently, multiple repeats (up to 28) were employed to establish reasonably narrow m/z standard deviations, which ranged from 1.47–2.97 for samples with ten or greater repeats. The corresponding 95% confidence intervals ranged from 0.54 to 1.71. Once mean experimental masses and variations were determined

Table 1
Amino-Terminal primary sequencing and MALDI-TOF determined masses for aging full length heme domain, *Bj*FixLH$_{140–270}$ and the −14C-terminal variant, *Bj*FixLH$_{140–256}$.

Age day[a]	ID.[b]	N-terminal sequence[c] experimental	N-term loss[d]	Mass[e] Expt'l	C-term loss[f]	Mass calculated	|Δ| mass[g]
0	*Bj*2	$_{(-1)}$M $_{140}$TRETHLRSILH	None	14,998.1	None	15,001.0	2.9
		$_{140}$TRETHLRSILH	None	14,868.4		14,869.8	1.4
7	*Bj*35	$_{140}$TRETHLRSILH	None	14,872.1	None	14,869.8	2.3
		$_{151}$TIPDAM(I)V	−11N				
		$_{145}$LRSILHTIPDA	−5N				
20	*Bj*31	$_{150}$HTIPDAMIVI	−10N		−14C		
		$_{151}$TIPDAMIVID	−11N				
		$_{147}$SIL_TIPDAMI	−7N				

[a] Number of 24 hour periods for which sterile room temperature aging was carried out.
[b] Laboratory reference identification of the specific *Bj*FixLH preparation.
[c] Parenthesis indicates ambiguous amino acid; assigned from sequence context.
[d] Amino-terminal truncation derived from primary sequencing, expressed as the number of deleted amino acids from the full length heme domain (FLHD), *Bj*FixLH$_{140–270}$, whose initial N-terminal sequence is TRETHLRSILHTIPDAMIVI-.
[e] Experimentally determined mass by MALDI-TOF mass spectrometry; one determination each.
[f] C-terminal truncation is −14 amino acids for *Bj*31 (*Bj*FixLH$_{140–256}$).
[g] Absolute difference between measured mass and calculated mass match with external calibration using Calmix-3 (Sigma).

Table 2
Low mass peak identification by mass matching between MALDI-TOF experimental masses and calculated masses for room temperature aging of *Bj*FixLH$_{140\text{-}256}$ (−14C).

Mass experimental[a]	S.D.[b]	95%[c]	N[d]	Mass match calculated	|Δ|[e] mass	Matched N-terminal losses[f]	Comments
13,309.96	1.61	1.00	10	13,309.22	0.74	–	+N terminal Met
13,177.44	2.30	1.30	12	13,178.02	0.58	–	−N terminal Met
12,440.51	1.90	1.03	13	12,440.20	0.31	−6N	−N terminal Met
12,284.72	1.79	0.67	27	12,284.02	0.70	−7N	−N terminal Met
11,971.63	1.47	0.54	28	11,970.62	1.01	−10N	−N terminal Met
11,833.92	2.10	0.80	27	11,833.48	0.44	−11N	−N terminal Met

[a] Measured mass for neutral compound.
[b] Standard deviation.
[c] 95% confidence interval.
[d] Number of determinations.
[e] Absolute difference between experimentally determined mass and calculated mass match.
[f] From the FLHD sequences: T(140)RETHLRSILHTIPD.

they were compared with all possible combinations of N-terminal and C-terminal amino acid deletions (calculated masses). For all three variant *Bj*FixLH used in this work, unique, closely matching masses between experimentally measured peaks and calculated protein sequences were found (Tables 2–4). This mass matching agreement was under 1.0 m/z for 15 of 19 peaks in this data set, and between 1.0–1.7 m/z for the remaining four peaks.

3. Results and discussion

3.1. Background

Previous observations of aging mass loss in nine different recombinant *Bj*FixLHs that were produced in three different laboratories were unexpected [1], as were similar mass losses that have been documented for *Sm*FixLH and *Ec*DosH (unpublished data). Subsequent, extensive aging experiments on *Bj*FixLH$_{(140\text{–}270)}$, our full-length recombinant heme domain, has revealed that observed mass losses were unaffected by many environmental factors, including presence or absence of: oxygen, heme group, protease inhibitor cocktails, AEBSF (or PMSF), light, ligands, sterilization, chemical reducing agents; whether glass or plastic containers were used, and differences in the heme iron ion oxidation state (unpublished data).

3.2. E. coli proteases as a possible cause of mass loss

The expression system for all proteins made in this laboratory employs *E. coli* strain BL21, a derivative of the K12 strain that has been engineered to lack the *Lon* and *OmpT* genes, that produce the most abundant proteases in *E. coli* K12. Aside from those deleted protease genes in strain BL21, the *E. coli* Protease Database (epd) [30] lists 71 other proteases and peptidases in *E. coli*. Many of these have only been inferred from genomics data mining, others are so poorly characterized that their substrates are to-date unknown. Proteases and peptidases are classified by their localization in the epd (cytoplasm, periplasm, cytoplasmic membrane, and periplasmic membrane). My working hypothesis is that only soluble proteases could persist through the isolation and purification procedures so that only cytoplasmic and periplasmic proteases/peptidase need be considered further.

There are 39 soluble proteases and peptidases in the cytoplasm of Strain K12. Two are the *Lon* proteases that are deleted in Strain BL21 and need not be further considered. Five of the remaining 37 are hypothetical, not yet isolated proteins, inferred from the genome. They have no determined functions, nor have their substrates been yet published. Twelve others have been noted but have as yet undetermined function and no identified substrates. Of the remaining twenty many are specific peptidases demonstrated to have substrates that are small peptides (described like “hydrolyzes tripeptides containing N-terminal Met, Leu or Phe”), or are proteases that cleave specific places, like “removes dipeptides from the C-termini of oligopeptides, broad specificity…” or “removes N-terminal methionine”. For those whose specificity is reported in the epd and for which at least representative substrates have been determined, no combination of them accounts for the observed cleavage points leading to the pattern of mass losses shown in Fig. 1 for the three terminal variants of *Bj*FixLH used in this work.

There are 21 soluble proteases and peptidases in the periplasm of Strain K12 and Strain BL21. One is a hypothetical protein, not yet isolated, with no proposed function and no proposed substrates. Eight have no determined function or proposed substrates. Of the remaining 12, six are reported to operate on carbohydrate–protein complexes that are involved in cell-wall formation. Two of the remaining six are identified as known serine proteases. The other four have specific activities, described like “hydrolyzes only long chain acyl thioesters (C12–C18 in length)”, or “heat shock protein with protease and chaperone properties”. Only the two serine proteases described above might have sufficient generality to produce the cleavage sites shown in Fig. 1 and these should be reliably disabled by our use of PMSF or AEBSF in addition to added protease cocktails in our isolation and purification procedures.

Table 3
Low mass peak identification by mass matching between MALDI-TOF experimental masses and calculated masses for room temperature aging of *Bj*FixLH$_{151\text{–}270}$ (−11N variant).

Mass experimental[a]	S.D.[b]	95%[c]	N[d]	Mass match calculated	|Δ|[e]	Matched C-terminal losses	Comments
13,656.29	1.81	0.74	23	13,656.50	0.21	–	Parent + N-Terminal Met
13,524.58	1.76	0.76	21	13,525.30	0.72	–	Parent − N-Terminal Met
12,889.66	1.91	0.72	27	12,888.61	1.05	−6C	+N-Terminal Met
12,757.29	1.99	0.79	24	12,757.41	0.12	−6C	−N-Terminal Met
12,688.00	1.70	1.87	8	12,686.33	1.69	−7C	−N-Terminal Met
12,558.79	1.92	1.42	7	12,558.19	0.60	−8C	−N-terminal Met

[a] Mass for neutral compound determined from mass spectrum.
[b] Standard deviation.
[c] 95% confidence interval.
[d] Number of determinations.
[e] Absolute difference between measured mass and calculated mass match.

Table 4
Low mass peak identification by mass matching between MALDI-TOF experimental masses and calculated masses for room temperature aging of *Bj*FixLH$_{140–264}$ (−6C variant).

Mass[a] experimental	S.D.[b]	95%[c]	N[d]	Mass match (calculated)	|Δ|[e]	Terminal deletions	Comments
14,102.66	1.96	0.80	23	14,101.96	0.70	−6C	Parent − N-Terminal Met
13,902.43	2.55	1.39	12	13,902.75	0.32	−8C	−N-Terminal Met
13,207.87	2.97	1.55	14	13,207.96	0.09	−7N −6C	−N-Terminal Met; Major peak
13,121.48	3.10	2.71	5	13,120.88	0.60	−8N −6C	−N-Terminal Met
13,007.63	2.76	1.71	10	13,007.79	0.09	−9N −6C	−N-Terminal Met; Major peak
12,893.95	9.85	9.65	4	12,894.56	0.61	−10N −6C	−N-Terminal Met
12 759.16	4.07	2.52	10	12,757.42	1.74	−11N −6C	−N-Terminal Met

[a] Mass for neutral compound determined from mass spectrum.
[b] Standard deviation.
[c] 95% confidence interval.
[d] Number of determinations.
[e] Absolute difference between measured mass and calculated mass match.

One can never prove the absence of a protease. However this laboratory's observations of similar mass losses for three species of heme-PAS domain proteins, made in three geographically remote laboratories [28] indicate that such a protease must be exceptionally resistant to standard protease inhibitors.

3.3. Description of the terminal deletion mutants used

A result of the previous work [1] was identification of a stable recombinant truncated *Bj*FixLH that did not exhibit mass loss while maintained over one year in solution. That protein, *Bj*FixLH$_{151–256}$, consists of 106 amino acids, compared to 131 amino acids present in the polypeptide of *Bj*FixLH$_{140–270}$. The degradation resistant truncated core PAS heme domain was thus missing 11 N-terminal (−11N) and 14 C-terminal (−14C) amino acids compared to the parent FLHD protein. That result could most logically be rationalized by hypothesizing that in heme PAS domains comprised of polypeptides longer than ~106 amino acids, the observed degradation was due to loss of amino acids from the proteins' termini. To test this hypothesis and to determine whether or not both termini were involved, aging experiments with mass and sequence monitoring were conducted. Two of the proteins used in this work were single terminus deletion variants (−11N or −14C) that each separately mimicked one of the deleted termini of the stable, 106 amino acid truncated core heme domain. The two variants were *Bj*FixLH$_{140–256}$ (−14C) and *Bj*FixLH$_{151–270}$ (−11N). For these two the goal was to determine the validity of the working hypothesis that, in each, aging mass loss only occurred in the fully intact terminus, not the truncated one. A third variant, *Bj*FixLH$_{140–264}$ (−6C), was used to determine whether the C-terminus deletion resulted in simplification of MALDI-TOF spectra (compared to the FLHD) sufficient for unambiguous peak assignments.

3.4. Integrity of variant BjFixLHs

Partial primary amino acid sequences of these proteins' N- and C-termini are shown in Fig. 1A, B, D, and F. The central core amino acid sequences are not explicitly shown, instead being represented by a dotted line connecting the explicit terminal sequences. Like their parent, *Bj*FixLH$_{140–270}$, the three variant proteins represented by sequences in Fig. 1B, D, and F all bind heme. This is demonstrated by the virtually identical ultraviolet–visible spectra displayed by the recombinant proteins *Bj*FixL$_{140–505}$, *Bj*FixLH$_{140–270}$, *Bj*FixLH$_{151–256}$, *Bj*FixLH$_{140–256}$, *Bj*FixLH$_{140–264}$, and *Bj*FixLH$_{151–270}$ in the Supplementary data (Appendix A). The spectral similarity shows that a native heme-bound core PAS structure is maintained by all recombinant PAS proteins used in this work, including the 106-amino acid PAS core heme domain. Whereas uv–visible spectra are little affected by changes in the termini, the effects of mass losses on ^{1}H hyperfine NMR spectra are substantial (unpublished data).

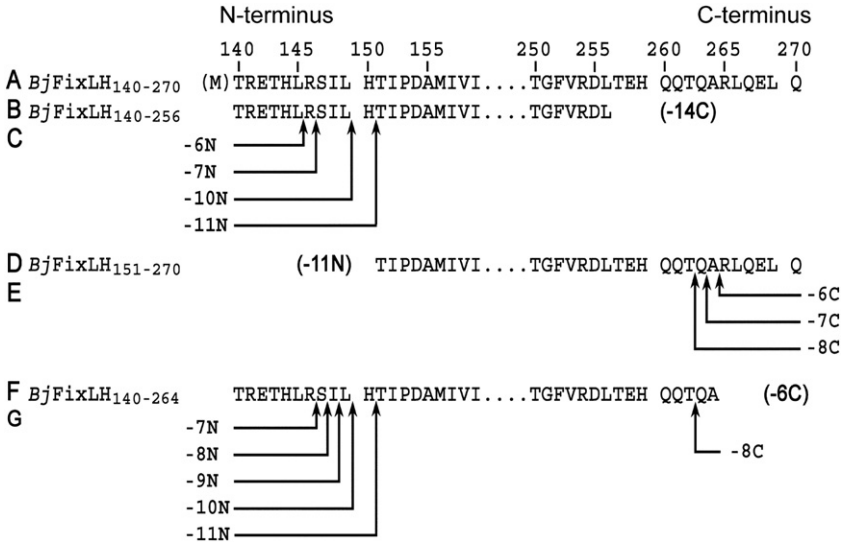

Fig. 1. Explicit N-terminal and C-terminal protein sequences of variant *Bj*FixLHs used in this work. A PAS core sequence consisting of 90 amino acids is represented by the dotted line. Line A represents the terminal sequence of the full length heme domain that is the parent sequence to the variants given in lines B (−14C), D (−11N) and F (−6C). Observed aging mass losses, and the cleavage sites for the individual variants are presented diagrammatically in C, E and G.

3.5. Primary sequencing for $BjFixLH_{140–270}$ (FLHD) and aged $BjFixLH_{140–256}$ (−14C)

Table 1 shows N-terminal primary sequencing results for a fresh preparation of *Bj*FixLH$_{140–270}$ (FLHD: *Bj*2) revealing the presence of two sequences that match those shown in Fig. 1A. The two sequences only differ in the presence of an N-terminal Met, a consequence of expression and the incomplete intracellular processing by methionine aminopeptidase in *E. coli*. Corresponding to this sequencing result, as noted before [1,28], and shown in Table 1, the experimental mass spectrum of the FLHD yields two measured parent protein masses separated by ~131 m/z, also indicating the variable Met labeling.

A separate *Bj*FixLH$_{140–270}$ preparation (*Bj*35) was subjected to amino-terminal primary sequencing after aging at room temperature for 7 days, giving the results presented in Table 1. In addition to the parent FLHD sequence, two truncated sequences were found, one corresponding to the loss of five N-terminal amino acids (−5N) and one corresponding to the loss of 11 N-terminal (−11N) amino acids. In a separate experiment the −14C truncated heme domain, *Bj*FixLH$_{140–256}$ (*Bj*31), was sequenced following 20 days of room temperature aging. No sequence corresponding to the non-degraded parent *Bj*FixLH$_{140–256}$ was identified, but three other sequences were, corresponding to losses of −7, −10 and −11, N-terminal amino acids, respectively. These primary sequencing results explicitly confirm one-half of the working hypothesis, namely that the N-terminus of the *Bj*FixLH polypeptide contributes to the observed spontaneous mass loss.

3.6. MALDI-TOF monitored aging of $BjFixLH_{140–256}$ (−14C)

Fig. 2 presents a stack plot of MALDI-TOF mass spectra for progressive aging. Day 0 represents the spectrum of the freshly prepared parent protein in the +1 ion region immediately following preparation for room temperature aging, including sterilization. Spectra obtained at days 4, 11 and 31 of aging are also presented. The Day 0 trace shows two parent protein peaks in the +1 ion range, corresponding to the presence (13,309.96 Da; Table 2) and absence (13,177.44 Da) of an N-terminal Met as discussed above. At subsequent days of aging, the spectra revealed progressive mass loss by the disappearance of the parent peaks and progressive increase in number and intensity of four resolved lower mass peaks. Those lower mass peaks occurred at 12,440.51, 12,284.72, 11,971.63 and 11,833.92 Da. Unambiguously identifying these peaks with a specific mass (hence primary sequence) required mass matching between the experimentally determined measured mass, the known starting sequence and calculated masses for the −14C variant, as described in Section 2. The quantitative results are presented in Table 2. In all cases the mass matches are quite close (Table 2) and they are unique, which indicates that none of the observed lower-mass peaks originate from further mass loss at the C-terminus, beyond the original −14 amino acid truncation. This agrees with the sequencing of a separate −14C preparation shown in Table 1 (*vide supra*). The amino acids lost from the N-terminus are diagrammatically shown in Fig. 1C.

3.7. MALDI-TOF monitored aging of $BjFixLH_{151–270}$ (−11N)

Fig. 3 shows a MALDI-TOF mass spectral stack plot (+1 ion region) consisting of the spectra of a freshly prepared sample of $BjFixLH_{151–270}$ (Day 0) immediately before starting room temperature aging (Day 0), and spectra from samples during the course of sterile aging taken for mass spectrometry at days 17 and 31. Peak masses and mass matching results are given in Table 3. As before, the Day 0 spectrum consists of two parent peaks (13,656.29; 13,524.58 Da), but in this case the peak corresponding to the heme domain without an N-terminal Met has the lower intensity. By aging Day 17 two peaks (12,889.66, 12757.29 Da) attain magnitudes comparable to the parent protein peaks and have masses corresponding to proteins that have lost six amino acids from the C-terminus (+/− N-terminal Met). At Day 31 spectra reveal two additional peaks, one identified as the −7C protein (12,688.00 Da) and the other identified as the −8C heme domain (12,558.79 Da). These amino acid losses are indicted in Fig. 1E. The mass matching data shown in Table 3 are again unique and indicate that during aging the N-terminal Met is also lost. The data indicate that losses of at least eight amino acids can occur from the C-terminus. In combination with the aging results described for the −14C variant

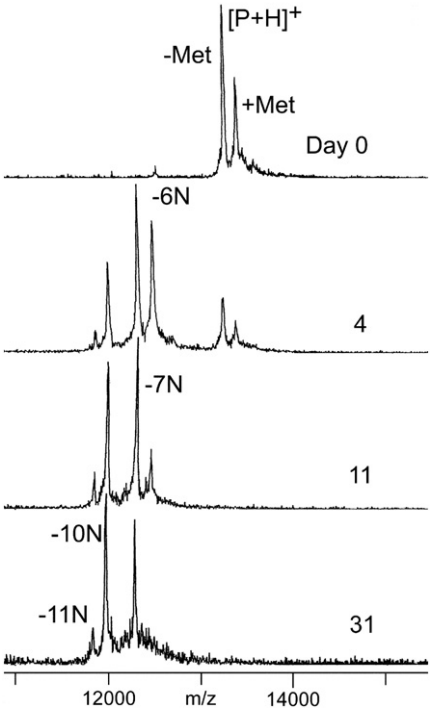

Fig. 2. Stack plot of MALDI-TOF mass spectra for *Bj*FixLH$_{140–256}$ (−14C) aging in the +1 ion region. The Day 0 spectrum is from the freshly prepared sample. Spectra at Days 4, 11 and 31 of the aging process are shown and reflect mass loss by the appearance of lower m/z peaks. Lower mass peaks corresponding to degraded heme domains are identified by their additional amino acid losses from the N-terminus.

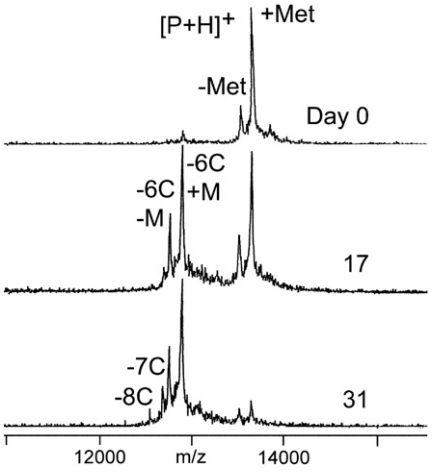

Fig. 3. MALDI-TOF mass spectra for *Bj*FixLH$_{151–270}$ (−11N) aging in the +1 ion region. The Day 0 spectrum is from the freshly prepared sample. Spectra at Days 17 and 31 of aging reflect mass loss by the appearance of lower m/z peaks. Lower mass peaks corresponding to degraded heme domains are identified by their additional amino acid losses from the C-terminus. M = N-terminal Met.

these data show that mass losses can occur from both full length termini, thereby validating the working hypothesis.

3.8. MALDI-TOF monitored aging of $BjFixLH_{140–264}$ (−6C)

This variant contains a complete N-terminal sequence (Fig. 1F) whereas its C-terminus is shortened by six amino acids, compared to $BjFixLH_{140–270}$ (Fig. 1A). The foregoing results have shown that aging degradation can proceed by amino acid losses in both termini, which we must therefore consider when interpreting aging results for $BjFixLH_{140–264}$ (−6C) and $BjFixLH_{140–270}$ (FLHD). MALDI-TOF spectra (unpublished) of $BjFixLH_{140–270}$ aging are quite complex; manifesting two broad envelopes with many overlapping peaks. The −6C variant was constructed to reduce such complexity by reducing the number of degraded N-terminal/degraded C-terminal combinations.

Spectral simplification was achieved compared to the FLHD (data not shown; see also ref [1] Fig. 2) as shown by the mass spectral stack plot of $BjFixLH_{140–264}$ aging at days 0, 17 and 31 in Fig. 4. Experimentally measured masses and the results of matching them with calculated masses to uniquely identify the extent of each N-terminus and C-terminus are shown in Table 4. Aging the −6C variant results in an increase in the number of lower mass peaks compared to each of the spectra for the −11N (Fig. 3) and −14C (Fig. 2) variants. It also demonstrates the broad envelope at m/z = 12,500–13,500 that is but one of two more intense and complex envelopes found in spectra of aging $BjFixLH_{140–270}$ (unpublished data).

Most major peaks identifiable in Fig. 4 are derived from N-terminal losses of 7–11 amino acids, paired with the C-terminal six amino acid deficit of the parent protein. Appearance of a major peak corresponding to the mass of a protein degraded by the loss of two additional C-terminal amino acids (i.e. −8C) shows that degradation can derive from both termini. Unreported in Fig. 4 and Table 4, but not unidentified in spectra are minor intensity peaks whose mass matches correspond to N-terminal amino acid losses of that −8C protein.

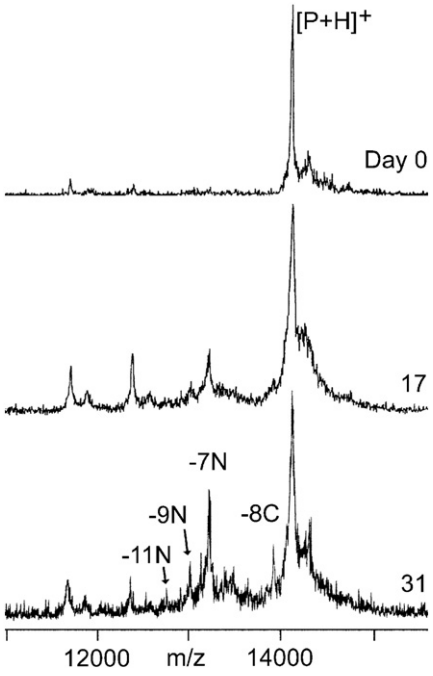

Fig. 4. MALDI-TOF mass spectra for $BjFixLH_{140–264}$ (−6C) aging in the +1 ion region. The Day 0 spectrum is from the freshly prepared sample. Spectra at Days 17 and 31 of aging reflect mass loss by the appearance of lower m/z peaks. Lower mass peaks corresponding to degraded heme domains are identified by their additional amino acid losses from the C-terminus.

4. Conclusions

The results have shown that mass loss from aging degradation can proceed by amino acid losses in both termini, which we must therefore consider when interpreting aging results for $BjFixLH_{140–270}$ (work in progress). The results presented confirm the working hypothesis formulated from the previous survey of recombinant *Bj*FixLH proteins' mass integrities [1]. It provides an operating paradigm upon which to base detailed analyses of the aging mass loss in the recombinant full length heme domain, $BjFixLH_{140–270}$, and its N–terminal His-tagged congener. The data presented above also provide degradation pattern data for terminal variants of $BjFixLH_{140–270}$, with cleavage points summarized in Fig. 1. A consistent concern in the ongoing studies of 15 different heme PAS proteins from three species has been whether the observed mass losses are simply due to one or more undetected and unarrested proteases. As discussed in Section 3.2 the cleavage sites are not, with currently available information, attributable to *E. coli*-specific protease activity. Such a protease would have to be unaffected by commercial protease inhibitor cocktails in addition to extra AEBSF (PMSF) that is present in our preparations from cell lysis through aging studies. It would also have to be present in the independent preparations of *Bj*FixLH and *Sm*FixLH from two laboratories that are geographically remote from this one.

Nomenclature

AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
PMSF	phenylmethylsulfonyl fluoride
Bj	*Bradyrhizobium japonicum*
Ec	*Escherichia coli*
epd	*E. coli* protease data base
Sm	*Sinorhizobium meliloti* (formerly *Rhizobium meliloti*)
FixL	two domain, heme + kinase protein
FixJ	phosphorylation substrate of FixL and response regulator, transcription factor
FixLJ	FixL + FixJ
FixLH	refers generally to any recombinant isolated heme domain
FLHD	our recombinant full length heme domain: $BjFixLH_{140–270}$;

Acknowledgments

I wish to thank Christine Suquet for making stocks of these proteins, for her patient instruction in the details of producing these proteins and for her many years of commitment to this project. The proteins were made prior to 2006 with funding provided by the National Institutes of Health, RO1 GM47645 (JDS).

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.jinorgbio.2011.01.015.

References

[1] J.D. Satterlee, C. Suquet, A. Bidwai, J. Erman, L. Schwall, R. Jimenez, Biochemistry 47 (2008) 1540–1553.
[2] D. Anthamatten, H. Hennecke, Mol. Gen. Genet. 225 (1991) 38–48.
[3] M.A. Sciotti, A. Chandon, H. Hennecke, H.-M. Fischer, J. Bacteriol. 185 (2003) 5639–5642.
[4] S. Mesa, F. Hauser, M. Friberg, E. Malaguti, H.-M. Fischer, H. Hennecke, J. Bacteriol. 190 (2008) 6568–6579.
[5] D. Nellen-Anthamatten, P. Rossi, O. Preisig, I. Kullik, M. Babst, H.-M. Fischer, H. Hennecke, J. Bacteriol. 180 (1998) 5251–5255.
[6] A.-M. Garnerone, C. Didier, M. Foussard, P. Boistard, J. Batut, J. Biol. Chem. 274 (1999) 32500–32506.
[7] C. Bobik, E. Meilhoc, J. Batut, J. Bacteriol. 188 (2006) 4890–4902.
[8] M.A. Gilles-Gonzalez, G.D. Ditta, D.R. Helinski, Nature 350 (1991) 170–172.
[9] E.K. Monson, M. Weinstein, G.S. Ditta, D.R. Helinski, Proc. Natl Acad. Sci. USA 89 (1992) 4280–4284.

[10] M.A. Gilles-Gonzalez, G. Gonzalez, M.F. Perutz, L. Kiger, M.C. Marden, C. Poyart, Biochemistry 33 (1994) 8067–8073.
[11] W. Gong, B. Hao, M.K. Chan, Proc. Natl Acad. Sci. USA 95 (1998) 15177–15182.
[12] K.R. Rodgers, Curr. Opin. Chem. Biol. 3 (1999) 158–167.
[13] M.K. Chan, Curr. Opin. Chem. Biol. 5 (2001) 216–222.
[14] D. Anthmatten, B. Scherb, H. Hennecke, J. Bacteriol. 174 (1992) 2111–2120.
[15] M.A. Gilles-Gonzales, G. Gonzalez, J. Biol. Chem. 268 (1993) 16293–16297.
[16] K.R. Rodgers, G.S. Lukat-Rodgers, J.A. Barron, Biochemistry 35 (1996) 9539–9548.
[17] J.R. Tuckerman, G. Gonzalez, E.M. Dioum, M.A. Gilles-Gonzalez, Biochemistry 41 (2002) 6170–6177.
[18] A. Tanaka, H. Nakamura, Y. Shiro, H. Fuji, Biochemistry 45 (2006) 2515–2523.
[19] C.M. Dunham, E.M. Dioum, J.R. Tuckerman, G. Gonzalez, W.G. Scott, M.A. Gilles-Gonzalez, Biochemistry 42 (2003) 7701–7708.
[20] J. Miksovska, C. Suquet, J.D. Satterlee, R.W. Larsen, Biochemistry 44 (2005) 10028–10036.
[21] R.W. Larsen, J. Miksovska, Coord. Chem. Rev. 251 (2007) 1101–1127, (specifically, pp 1116–1118).
[22] W. Gong, B. Hao, M.K. Chan, Biochemistry 39 (2000) 3955–3962.
[23] B. Hao, C. Isaza, J. Arndt, M. Soltis, M.K. Chan, Biochemistry 41 (2002) 12952–12958.
[24] J. Key, K. Moffat, Biochemistry 44 (2005) 4627–4635.
[25] J. Key, V. Srajer, R. Pahl, K. Moffat, Biochemistry 46 (2007) 4706–4715.
[26] R.A. Ayers, K. Moffat, Biochemistry 47 (2008) 12078–12086.
[27] H. Miyatake, J. Mukai, S.-Y. Park, S.-I. Adachi, L. Tamura, H. Nakamura, K. Nakamura, T. Tsuchiya, T. Iizuka, Y. Shiro, J. Mol. Biol. 301 (2000) 415–431.
[28] C. Suquet, M. Savenkova, J.D. Satterlee, Protein Expression Purif. 42 (2005) 182–193.
[29] J.D. Satterlee, C.M. Suquet, M. Savenkova, C. Lian, ACS Symp. Ser. 858 (2003) 244–257.
[30] www.cf.ac.uk/biosi/staffinfo/ehrmann/tools/proteases.index.html, last accessed November 19, 2010.