S11-022

Mass Spectrometry of the Subunits of Highly Active Cytochrome b_6f Complex from Spinach and the Cyanobacterium, *M. laminosus*: Identification of petA,B,C,D,G,L,M,N Subunits in Both Complexes and FNR in the Spinach b_6f Complex.

J.P. Whitelegge[1], H. Zhang[2], W.A. Cramer[2]
1. The Pasarow Mass Spectrometry Laboratory, Departments of Chemistry & Biochemistry, Psychiatry & Biobehavioural Sciences and the Neuropsychiatric Institute, University of California Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095, USA.
2. Dept. of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA.
Tel. 310 206 7886, Fax. 310 206 2161, jpw@chem.ucla.edu

Keywords: chloroplast proteome, thylakoid proteomics, cytochrome b_6f complex, ferredoxin:$NADP^+$ oxidoreductase, mass spectrometry.

Introduction

Mass spectrometry (MS) has revolutionized the biological sciences since the development of matrix assisted laser desorption ionization (MALDI) and electrospray-ionization (ESI) in the late eighties. Biological macromolecules are now mass measured with great accuracy and highly resolved spectra reveal subtle molecular heterogeneity. For intact proteins, ESI-MS provides ultimate accuracy (within 0.01 % error) and resolution while MALDI is more sensitive and more tolerant of extreme heterogeneity and complex mixtures. Thus a mass spectrum of an intact protein defines the native covalent form of a gene's product and associated heterogeneity (Whitelegge et al, 1998). While intrinsic membrane proteins have been traditionally regarded as problematic, a suite of techniques has been developed allowing ESI-MS of numerous examples with up to fifteen transmembrane helices (Whitelegge et al, 1998; 1999; le Coutre et al, 2000; Turk et al, 2000). The entire proteome is now accessible to mass spectrometry, though proteomics has largely been fueled by the ability of mass spectrometry to *identify* proteins based upon mass and sequence analysis of small peptides derived from the parent protein (Whitelegge & le Coutre, 2001). Previously ESI and MALDI were applied to polypeptides of photosystem 2 (Gómez et al, 1999; Whitelegge et al, 1997; 1998; 1999) including the larger subunits D1, D2, CP43 and CP47 from higher plants.

In this study, ESI-MS was used to characterize the cytochrome b_6f complex from spinach and a cyanobacterium *Mastigocladus laminosus* The products of petA,B,C,D,G,L,M,N are defined with respect to their intact mass and heterogeneity and a ninth subunit, associated with the higher plant complex only, was identified as ferredoxin-NADP oxidoreductase (FNR) via proteomics of peptides derived from it. Sequences of the smaller spinach subunits were determined by mass spectrometry revealing a number of modifications not apparent from translations of current genomic database entries.

Materials and Methods

Samples of isolated cytochrome b_6f complex were analyzed by liquid chromatography with mass spectrometry and fraction collection (LCMS+). Protein (200 µg) was precipitated with chloroform/methanol and dissolved in 60 % formic acid (100 µL) prior to loading onto a reverse-phase column (PLRP/S, 300 A, 5 µm, 2 x 150 mm; Polymer Lab.s) previously equilibrated in 95 % A, 5 % B (A, 0.1 % TFA in water; B, 0.05 % TFA

in acetonitrile/isopropanol, 1:1) at 40 ˚C. A linear gradient through 100 % B was developed over 60 minutes and the eluent passed through a UV detector before being split with half the flow directed to an electrospray-ionization source and the remainder to a fraction collector. Electrospray-ionization mass spectrometry (ESI-MS) was performed as described (Whitelegge et al, 1998). Data was processed using Hypermass or Multiview software (Applied Biosystems). Fractions collected concomitant with ESI-MS were treated with CNBr (100 mg/mL in 10 % formic acid) and the resulting peptides mass measured by ESI-MS. The sequence of the small subunits of the spinach complex were obtained using size-exclusion chromatography with tandem mass spectrometry (SEC-MSMS) and interpreted with the aid of homologous sequences from the database. SEC was performed in chloroform/methanol/1 % aqueous formic acid (4/4/1, v/v; Whitelegge et al, 1999) and SEC-MS used for MS of PetN,L.

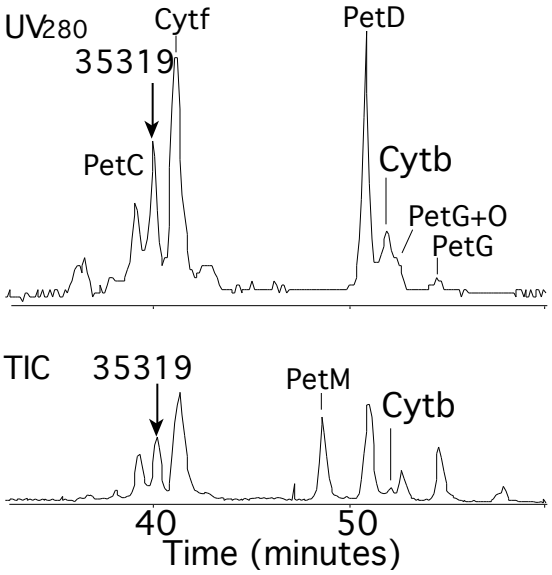

Figure 1. LCMS+ of spinach cytochrome *b*$_6$*f* subunits. A reverse-phase separation is eluted directly to an electrospray-ionization mass spectrometer with a split line allowing simultaneous collection of fractions. Cytochrome b elutes with low efficiency just before singly oxidized PetG.

Results and Discussion

The subunits of cytochrome *b*$_6$*f* that precipitated in chloroform/methanol were separated by reverse-phase chromatography with ESI-MS and fraction collection (LCMS+; Figure 1). PetN and PetL are lipoproteins, partitioning into the chloroform phase of the precipitation, and were thus not recovered in this experiment. The UV elution profile is compared with the total ion chromatogram (TIC; a measure of total mass spectrometer response with time) demonstrating resolution of seven subunits of the spinach complex. The identities of each peak were assigned by mass except for the 35319 Da peak which was not a known component of the complex but potentially an apoprecytochrome *f* or some other component. Proteomics of a fraction collected as the 35319 Da protein eluted identified the molecule as ferredoxin-NADP oxidoreductase (FNR; see S11-002) defining the ninth component of chloroplast cytochrome

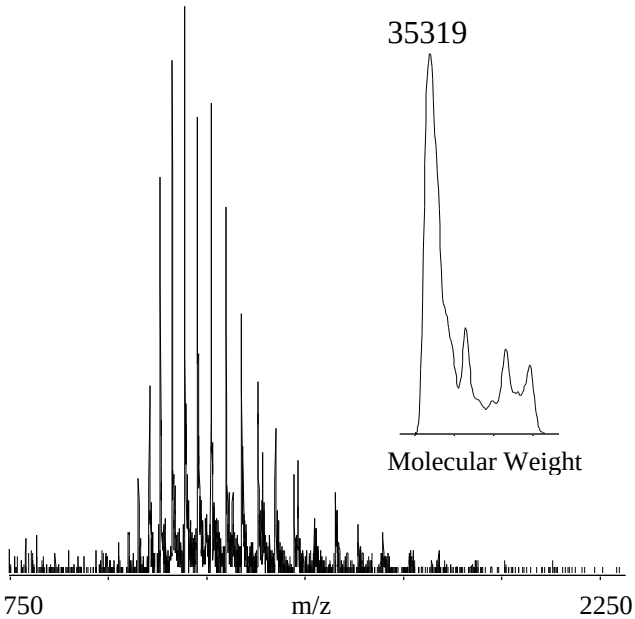

Figure 2. ESI-MSof the newly discovered ninth subunit of chloroplast cytochrome *b*$_6$*f* complex.

b_6f complex. The ESI mass spectrum of the 35kDa subunit is shown in Figure 2. ESI results in multiple charging of the protein ions resulting in a cluster of signals of varying m/z, the most abundant of which here carries 37 protons (m/z = $[M+37H]^+/37$). A zero-charge molecular weight spectrum is shown inset emphasizing the limited heterogeneity associated with the 35 kDa subunit. The nature of the minor components is unknown.

All of the subunits were characterized in a similar fashion, including cytochrome *b* which has eight transmembrane helices and elutes with sub-stoichiometric efficiency during chromatography under these conditions. The masses of the nine chloroplast and eight cyanobacterial subunits measured with ESI-MS are shown in Table 1.

Table 1. Masses of spinach and *M. laminosus* cyt.*b_6f* complex subunits.

Protein	Measured Mass	Calculated Mass	Modifications
S. oleracea	(Da)	(Da)	
PetC	18938	18938.6	69on, + 1 disulfide
FNR	35319	35313.7	
Cytf	31940	32035.8	36on, + heme
Cytb	24884	24782.1	+ heme
PetD	17313	17313.6	Minus Met1
PetN	3198	3197.84	N-formylation
PetL	3478	3478.24	formyl, **S2->F**
PetM	3973	3972.70	Modified *A.thaliana*
PetG	4198	4198.03	N-formylation
M. laminosus			
PetC	19295	19202.8	
Cytf	32273	32269.4	+ heme
Cytb	24712	24884.2	+ heme
PetD	17528	17521.9	
PetN	3304		
PetL	3530		
PetM	3841		
PetG	4058		

In some cases the measured masses agree very well with masses calculated based upon annotated database entries (SwissProt). In spinach PetC, PetD are in full agreement with current entries while greater uncertainties surround Cyt.*f* and Cyt.*b* and the cyanobacterial subunits which are less well characterized. Clearly further analysis is necessary before the primary structures of many of these proteins can be confirmed. The measured mass of the subunit now known to be FNR is not within the envelope of 0.01 % error but subsequent measurements have provided values closer to 35314 Da in support of the primary structure in SwissProt.

Toward a complete description of the subunits of the cytochrome *b_6f* complex, the small subunits of the spinach complex were subjected to tandem mass spectrometry (MSMS) to generate sequence data. A specific ion is selected and then fragmented by collision-activated dissociation (CAD) that typically breaks polypeptides at the peptide bond generating both N-terminal (*b*-ions) and C-terminal (*y*-ions). Existing sequence from the database is used to aid interpretation of MSMS spectra. The PetN sequence from the recently completed spinach chloroplast genome sequence (Schmitz-Linneweber et al, 2001) was formylated at the initiating methionine producing a perfect match to the measured mass (Table 1) and MSMS data that included enough b and y ions to confirm the internal sequence (MSMS data not shown). PetG was assigned by N-formylation of

the sequence translated from the same genome. In the case of PetL it was necessary to modify the published sequence S2 -> F to match measured mass and MSMS data. Whether this discrepancy is due to a DNA sequencing error or an RNA editing event is unclear. The nuclear-coded PetM was assigned by homology to the *A.thaliana* gene

Table 2. Sequences of small subunits by tandem mass spectrometry (MSMS).

Subunit	Sequence	MW
PetG	formylMIEVFLFGIVLGLIPITLAGLFVTAYLQYRRGDQLDL	4198.03
NP054954	MIEVFLFGIVLGLIPITLAGLFVTAYLQYRRGDQLDL	
PetM	NAVGEIFKIAAIMNALTLVGVAVGFVLLRIEATVEEAE	3972.70
A.thaliana	NAVGEIFKIAAIMNALTLVGVAVGFVLLRIE**TS**VEEAE**AE**	
PetL	formylMFTLTSYFGFLLAALTITSALFIGLNKIRLI	3478.24
NP054953	M**S**TLTSYFGFLLAALTITSALFIGLNKIRLI	
PetN	formylMDIVSLAWAALMVVFTFSLSLVVWGRSGL	3197.84
NP054925	MDIVSLAWAALMVVFTFSLSLVVWGRSGL	

with alteration of positions 32 & 33, removal of two C-terminal amino acids, and identical removal of leader peptide from the N-terminus. Clearly, complete annotation of genomic data requires detailed analysis of primary structure by mass spectrometry. The ability to monitor subtle covalent modifications, including oxidation or phosphorylation, that modulate physiological changes is emphasized (see Whitelegge, 1996; 1998).

Acknowledgements

Support from NIH, W.M. Keck Foundation and the Pasarow family is thanked.

References

Gómez SM, Park JJ, Zhu J, Whitelegge JP, Thornber JP (1999) In '*Photosynthesis: Mechanisms and Effects; Proceedings of the 11th International Congress on Photosynthesis*' (G. Garab, ed.), Kluwer Academic Publishers, vol. **I** 353-356.

le Coutre J, Whitelegge JP, Gross A, Turk E, Wright EM, Kaback HR, Faull KF (2000) *Biochemistry* **39**, 4237-4242.

Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, MacheR (2001) *Plant Mol Biol* **45**, 307-315.

Turk E, Kim O, le Coutre J, Whitelegge JP, Eskandari S, Lam JT, Kreman M, Zampighi G, Faull KF, Wright EM (2000) *Journal of Biological Chemistry* **275**, 25711-25716.

Whitelegge, JP (1996) In *'The Handbook of Photosynthesis*', (M. Pessarakli, ed.), Marcell Dekker, New York, pp. 241-256.

Whitelegge, JP (1998) In *The Handbook of Plant and Crop Stress*, (M Pessarakli, ed.), Marcell Dekker, New York. pp 555-568.

Whitelegge JP, le Coutre J (2001) *American Journal of Pharmacogenomics* **1**, 29-35.

Whitelegge JP, Gómez SM, Stevens RL, Mastrogiacomo A, Gundersen C, Faull KF (1997) *FASEB J* **11(9)**, A1108, 1357.

Whitelegge JP, Gundersen C, Faull KF (1998) *Protein Science* **7**, 1423-1430.

Whitelegge JP, Faull KF, Gundersen C, Gómez SM (1999) In '*Photosynthesis: Mechanisms and Effects; Proceedings of the 11th International Congress on Photosynthesis*' (G. Garab, ed.), Kluwer Academic Publishers, vol. **V** 4381-4384.

Whitelegge JP, le Coutre J, Lee JC, Engel CK, Privé GG, Faull KF, Kaback HR (1999) *Proceedings of the National Acadamy of Sciences U.S.A.* **96**, 10695-10698.