Mass Spectrometry Guided In Situ Proteolysis to Obtain Crystals for X-Ray Structure Determination

Tarun Gheyi, Logan Rodgers, Richard Romero, J. Michael Sauder, and Stephen K. Burley

New York SGX Research Center for Structural Genomics, Structural Biology, Eli Lilly and Company, Lilly Biotechnology Center, San Diego, California, USA

A strategy for increasing the efficiency of protein crystallization/structure determination with mass spectrometry has been developed. This approach combines insights from limited proteolysis/mass spectrometry and crystallization via in situ proteolysis. The procedure seeks to identify protease-resistant polypeptide chain segments from purified proteins on the time-scale of crystal formation, and subsequently crystallizing the target protein in the presence of the optimal protease at the right relative concentration. We report our experience with 10 proteins of unknown structure, two of which yielded high-resolution X-ray structures. The advantage of this approach comes from its ability to select only those structure determination candidates that are likely to benefit from application of in situ proteolysis, using conditions most likely to result in formation of a stable proteolytic digestion product suitable for crystallization. (J Am Soc Mass Spectrom 2010, 21, 1795–1801) © 2010 American Society for Mass Spectrometry

Knowledge of protein's three-dimensional (3-D) structure is often critical for understanding function. Various techniques, including X-ray crystallography, nuclear magnetic resonance (NMR), mass spectrometry (MS), circular dichroism (CD), Fourier-transform (FT) infrared (IR), Raman, ultraviolet (UV)-visible absorption and fluorescence spectroscopies, are used for protein analyses [1, 2]. Among these methods, X-ray crystallography has surpassed all other techniques in revealing structural information at atomic resolution for macromolecules, thereby contributing the majority of protein structures to the Protein Data Bank (www.pdb.org) [3]. Although productivity in this field continues to accelerate, overall failure rates in determining the structure of a given target remain high [4–6]. Success depends both on the stability and solubility of the target protein and on surveying appropriate crystallization conditions. As a result, protein engineering has emerged as an important means of improving protein physicochemical properties, which can render them more amenable to crystallization [7].

Since the development of MALDI and ESI modes of ionization for macromolecules [8–10], MS has become an essential tool for protein crystallographers. Within structural proteomics, MS is now routinely used for sample quality control, crystal content verification, analysis of selenomethionine incorporation and post-translational modifications, and identification of oligomerization states [11]. MS is also frequently used to identify stable domains within larger proteins [12–17]. The power of limited proteolysis/mass spectrometry (LPMS) to identify the N- and C-termini of stable domains is well documented and is frequently used to support design of expression constructs that produce proteins suitable for crystallization. Recently, in situ proteolysis has emerged as a productive strategy for modifying proteins that otherwise prove recalcitrant to crystallization [4, 18]. In favorable cases, one or both of these established proteolysis-based approaches can improve the chances of obtaining diffraction quality crystals. However, there are many proteins that do not succumb readily to either strategy, and there is a need for a more refined approach to the problem.

Herein, we describe a simple, generic protocol for identifying the optimum protease:protein ratios to generate stable protein domains on a time scale comparable to the duration of crystallization trials. The procedure is cost-effective and robust. MS provides a rapid/straightforward read out of the extent of polypeptide chain cleavage and the long-term stability of protein domains generated by limited proteolysis. The utility of this approach was tested by applying it to a representative set of 10 protein targets of unknown structure. Our protocol yielded diffraction quality crystals in four cases, leading to determination of two novel high-resolution structures. Importantly, it also identified four targets as being unsuitable for a proteolytic approach to protein engineering.

Address reprint requests to Dr. T. Gheyi, New York SGX Research Center for Structural Genomics, Structural Biology, Eli Lilly and Company, Lilly Biotechnology Center, 10300 Campus Point Drive, Suite 200, San Diego, CA 92121, USA. E-mail: Gheyi_Tarun@lilly.com

1044-0305/10/$32.00
doi:10.1016/j.jasms.2010.06.015

Published online July 7, 2010
Received April 23, 2010
Revised June 21, 2010
Accepted June 22, 2010

Materials and Methods

Cloning, Expression, and Purification of Proteins

Genes encoding the 10 protein targets of interest were cloned from genomic DNA or from codon-optimized synthetic templates. The desired amino acid boundaries were PCR amplified and the purified PCR products were subsequently TOPO (Invitrogen, Carlsbad, USA) cloned into pSGX3, a derivative of pET26b(+), that expresses protein with a noncleavable C-terminal hexa histidine tag. The resulting plasmids were transformed into BL21(DE3)-Condon+ RIL (Invitrogen) cells for overexpression. Se-Met protein production was carried out at 22 °C in 1L of high yield (HY) media (Orion Enterprises, Inc, Northbrook, IL, USA) containing 50 μg/mL of kanamycin and 35 μg/mL of chloramphenicol. Protein expression was induced by addition of 0.4 mM IPTG. Cells were harvested after 21 h by centrifugation at 4 °C.

For purification, the *E. coli* cell pellets were resuspended in cold buffer [20 mM Tris HCl pH 8.0, 500 mM NaCl, 25 mM imidazole, and 0.1% (vol/vol) Tween20] and were lysed via sonication. Cellular debris was removed by centrifugation at 4 °C and the decanted supernatant from each sample was applied to a 5 mL HisTrap-HP column (GE Health Care, Piscataway, NJ, USA) charged with nickel and pre-equilibrated with 20 mM Tris HCl pH 8.0, 500 mM NaCl, 10% (vol/vol) glycerol, and 25 mM imidazole. Bound samples were washed with 5 column volumes (CV) of 20 mM Tris HCl pH 8.0, 500 mM NaCl, 10% (vol/vol) glycerol, and 40 mM imidazole and subsequently eluted with 2 CVs of the same buffer, with an imidazole concentration of 250 mM. Eluted proteins were further purified using a 120 mL Superdex 200 size-exclusion column pre-equilibrated with 10 mM HEPES pH 7.5, 150 mM NaCl, 10% (vol/vol) glycerol, and 5 mM DTT (protein storage buffer). SDS-PAGE analysis showed greater than 95% purity and protein fractions corresponding to the symmetric portion of the size exclusion chromatography profile were pooled for concentration using AMICON (Millipore, Billerica, MA, USA) spin filters. Concentrated protein aliquots were flash frozen using liquid nitrogen and stored at −80 °C.

Protein Characterization

Matrix assisted laser desorption/ionization (MALDI)-MS (Voyager, DE-RP; Applied Biosystems, Foster City, CA, USA) and a liquid chromatography (LC)-electrospray ionization (ESI) mass spectrometer equipped with a quadrupole mass analyzer (API 150EX; Applied Biosystems) were used to determine the molecular masses of all protein samples. This information was required to assess both purity and chemical homogeneity, and to compare measured molecular masses to the calculated molecular masses derived from the expected protein sequence. In situ proteolysis mass spectrometry experiments were performed on protein samples with greater than 95% purity.

Proteolytic Digestion and Mass Spectrometry

Ten μL of each target protein (~2 mg/mL concentration) was independently incubated with three proteases (Proteinase K, Subtilisin and Chymotrypsin; Roche, Indianapolis, IN, USA) at room temperature over 9 days (which is comparable to the duration of crystallization trials). Four different protease-to-target ratios were used, including 1:5000, 1:10,000, 1:50,000, and 1:100000. Proteolytic reactions were conducted in protein storage buffer. Stock solutions of proteases were also prepared in protein storage buffer. Reactions were quenched at predetermined intervals by aliquoting 1 μL of reaction mixture into 1 μL of 5% TFA. Extent of proteolysis was monitored by mass analysis using MALDI-MS. A modified "thin-layer" method [19] was used for sample preparation. To improve the reproducibility of the sample analysis, saturated α-cyano-4-hydroxy-cinnamic acid (HCCA) matrix solution was prepared by dissolution into a mixture of formic acid, water, and isopropanol (a ratio of 3:1:2, respectively). Matrix solution was warmed in a boiling water bath for 2 min and maintained at 40 °C during sample preparation; 0.5 μL of the protein digest was mixed with 9.5 μL of warm matrix solution, and 0.2 μL of the final mixture was used for spotting on pre-layered MALDI sample plates.

MALDI-MS measurements were performed using a Voyager DE-RP mass spectrometer equipped with a pulsed N_2 laser (λ = 337 nm, pulse width 4 ns, frequency 20 Hz). MALDI mass spectra were obtained in positive linear mode by averaging the results of 300 individual laser shots. Thirty-six protease:protein digests (derived from three protein samples) were spotted on a MALDI-MS sample plate around a standard (for calibration). Mass calibration was carried out with carbonic anhydrase as an external standard (MW 29027 Da) using singly and doubly charged ionic species. Mass error $>$ 0.1% was expected due to spatial distribution and ignored because of subsequent ESI-MS analysis (mass error $<$ 0.01%); this procedure permitted more rapid analysis. MALDI-MS analysis was performed on alternate days (up to 9 d) to monitor the extent of protease digestion. No further mass analyses were performed on protein samples showing complete degradation and/or precipitation. If necessary, sites of proteolytic cleavage were identified by more accurate mass measurement using LC-ESI-MS. These measurements were conducted at an ion spray voltage of 5500 V, source temperature of 300 °C, a focusing potential of 250 V, and a scan range of 1100 to 1900 *m/z*. Protein molecular masses were estimated using BioAnalyst ver. 1.4.

Crystallization and X-Ray Data Collection

Protein samples were mixed with the appropriate amount of protease and immediately subjected to crystallization screening (0.3 μL protein + 0.3 μL reservoir solution) with the Classics, Classics II, and PEGs kits (Qiagen, Valencia, CA, USA) using a Phoenix Liquid Handling System (Art Robins, Sunnyvale, CA, USA) via sitting drop vapor diffusion at 21 °C. On first examination of these initial crystallization experiments 24 h later, there were significant numbers of crystalline hits compared with the results of control experiments that excluded proteases. Crystals obtained at various conditions were flash frozen by direct immersion in liquid nitrogen after cryo-protection via addition of 20% (vol/vol) ethylene glycol. Diffraction data were recorded using the LRL-CAT 31-ID beamline at the Advanced Photon Source (APS), and processed with MOSFLM [20] and SCALA [21] (Collaborative Computing Project Number 4, 1994).

Results and Discussion

The observations described in this paper and the collective experiences within structural biology as a discipline underscore the challenges of working with full length proteins. In favorable cases, proteins such as enzymes secreted by the exocrine pancreas [22] that have evolved to be biochemically active at extremes of pH, ionic strength, presence of proteases, etc. tend to have compact globular structures. Such enzymes were among the first proteins to be successfully studied by X-ray crystallography. Frequently, their inherently stable tertiary structures (Figure 1a) support crystallization without the need for protein engineering. At the NYSGXRC, we have had similar experiences with many members of the enolase superfamily. These proteins proved to be particularly well suited to crystallization with standard crystallization protocols [23] resulting in a success rate of 57% in going from homogenous, soluble purified protein to a 3D structure (http://pepcdb.pdb.org). Many proteins of interest to structural biologists, particularly those from eukaryotes, do not behave in this fashion. Figure 1b to e illustrate some of the more challenging target categories for which protein engineering is often required to identify and expunge flexible polypeptide chain segments that can impede crystal formation.

To overcome these challenges, we have combined both LPMS and in situ proteolysis strategies to identify the optimal protease concentration yielding a stably truncated form of the target protein. Proteolysis is

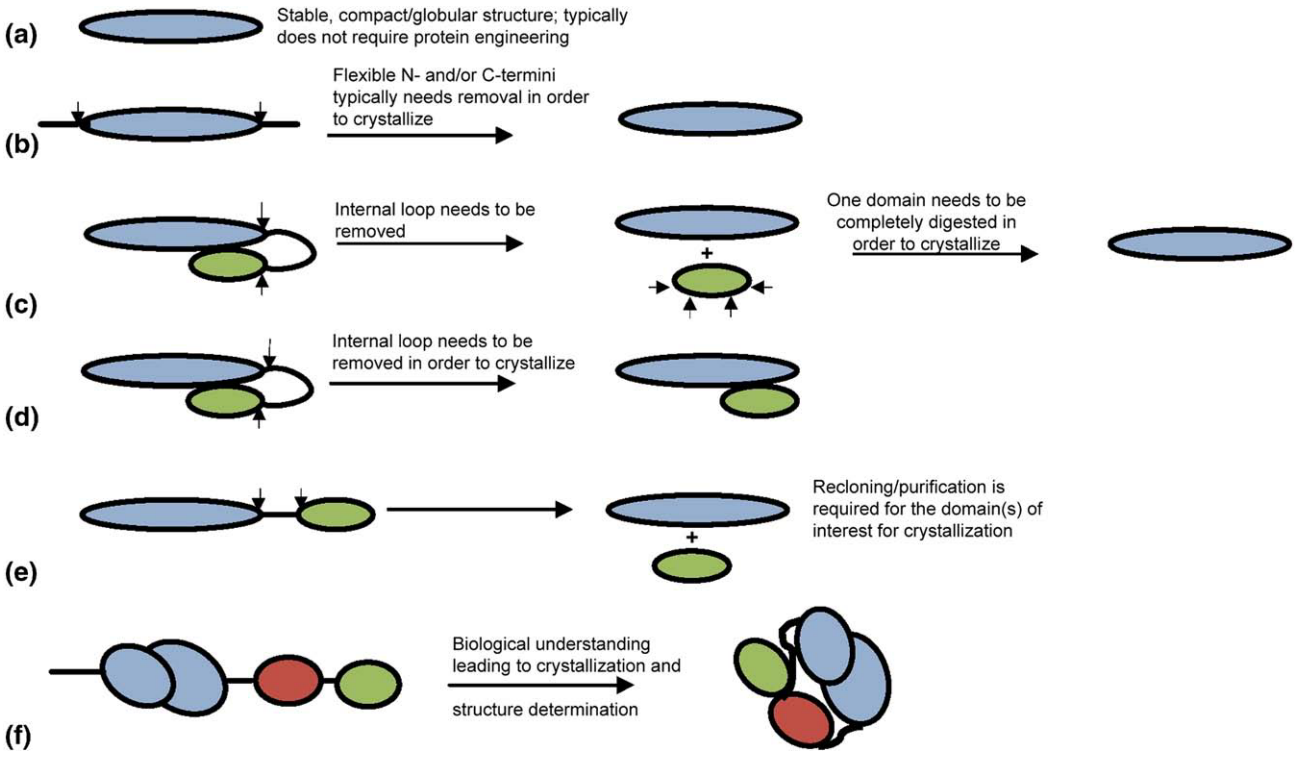

Figure 1. Schematic view of protein characteristics that complicate crystal formation. (**a**) Compact structure that readily forms crystals; (**b**) flexible N- and/or C-termini that need to be removed for crystal formation; (**c**) extra domain and disordered loop that must be removed to form crystals of the domain of interest; (**d**) disordered loop that must be removed to crystallize two domains; (**e**) interacting domains that must be re-expressed or re-purified after excising a disordered loop; (**f**) special case of a multi-domain protein, in which understanding the biology led to preparation of a modified protein that formed a compact arrangement of domains amenable to X-ray analysis [26].

intended to remove flexible segment(s) from the polypeptide chain; it may be at the N-terminus and/or the C-terminus or within solvent-exposed loop(s) separating globular regions of the polypeptide chain. MS identified combinations of protease(s) and protease/concentration(s) can be used for in situ proteolysis crystallization trials to increase the likelihood of successful structure determination. We describe our experience using in situ proteolysis guided by prior MS analyses to provide a rational means of identifying a target compatible protease and its optimal concentration for subsequent, directed in situ proteolysis crystallization trials.

As part of the technology development activities undertaken by the NIH-funded NYSGXRC structural genomics consortium (New York SGX Research Center for Structural Genomics, www.nysgxrc.org), 10 representative protein targets of unknown structure were selected for MS-in situ proteolysis experiments (Table 1). Although all of these proteins could be purified to homogeneity, none yielded crystals using our standard crystallization protocol [23]. Each of these 10 proteins

Table 1. Mass Spectrometry aided in situ proteolysis of proteins that had previously failed to produce diffraction quality crystals

Protein ID	Chymotrypsin				Proteinase K				Subtilisin			
15547	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SH → DQC → PDB	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SH → DQC	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SH → DQC	1/5 × 10^3 → CD
14771	1/10^5 → CD	1/5 × 10^4 → CD	1/10^4 → CD	1/5 × 10^3 → CD	1/10^5 → CD	1/5 × 10^4 → CD	1/10^4 → CD	1/5 × 10^3 → CD	1/10^5 → CD	1/5 × 10^4 → CD	1/10^4 → CD	1/5 × 10^3 → CD
15101	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SI	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SH → NC	1/5 × 10^3 → SH → NC	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SI
14639	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → ND	1/5 × 10^3 → ND	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SI	1/5 × 10^3 → SI	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SI	1/5 × 10^3 → SI
15527	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SI	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SH → NC	1/10^5 → ND	1/5 × 10^4 → ND	1/10^4 → SI	1/5 × 10^3 → SH → NC
15546	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SI	1/5 × 10^3 → SI	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SH → DQC → PDB	1/5 × 10^3 → SI	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → SI	1/5 × 10^3 → SH → DQC
16279	1/10^5 → SI	1/5 × 10^4 → SI	1/10^4 → CD	1/5 × 10^3 → CD	1/10^5 → SH → DND	1/5 × 10^4 → SH → DND	1/10^4 → CD	1/5 × 10^3 → CD	1/10^5 → SH → DND	1/5 × 10^4 → SI	1/10^4 → CD	1/5 × 10^3 → CD

A total of 12 protein:protease combinations gave stable homogenous forms of digested product leading to two high resolution novel structures. Three protein samples, which precipitated within 24 h were not used for further analysis and are not included in this table.
SH = stable homogenous form of digested protein; SI = heterogenous mixture of digested and undigested protein; CD = complete digestion; DQC = diffraction quality crystals; PDB = structure deposition; NC = no crystals; DND = did not diffract.
Targets used in this study: PID 15547, NYSGXRC-11030h, P24242: 3-335; PID 14771, NYSGXRC-9269c, gi 44624469: 2-394; PID 15101, NYSGXRC-10086f, Q71ZA7: 2-302/T189S; PID 14639, NYSGXRC-10057f, Q7WUJ6: 2-399/I309M; PID 15527, NYSGXRC-11020w, Q6G3E4: 23-374; PID 15546, NYSGXRC-11029y, P0ACZ4: 2-204; PID 16279, NYSGXRC-11030w, O06987: 50-316.

was independently digested using three different enzymes at four different concentrations. In all, 12 different combinations (protease × protease:protein concentration ratios) were surveyed for each protein sample. Cleavage products were analyzed by MALDI-MS to identify the best protease and protease:protein ratio for use in subsequent, directed in situ proteolysis crystallization trials. A schematic representation of the protocol is provided in Figure 2. MALDI-MS is the method of choice for molecular mass measurements because it provides rapid, reproducible, and easy-to-read mass spectra. Moreover, this technique is compatible with various salts/buffers used routinely in protein purification.

Figure 3a shows the MALDI mass spectrum of one of our purified protein targets (an *E. coli* transcriptional repressor, Genbank NP_417194, PSI TargetID "NYSGXRC-11030h"). The predicted mass of the protein plus the affinity purification tag (NYSGXRC proteinID 15547) is 38362 Da (346 residues, including a nonremovable C-terminal His tag). Within experimental error, the observed mass matches the predicted mass derived from a sequence-verified expression plasmid. Figure 3b and c show the MALDI mass spectra after digesting the same sample with chymotrypsin at 1:5000 (protease:protein) ratio for one and nine days, respectively. A stable domain of ~32 kDa molecular mass was produced after 1 day of digestion (no further digestion was observed over the following 8 d). Accurate mass measurement using ESI-MS identified the precise N- and C-termini of the chymotrypsin cleavage product. In this particular case, the calculated mass was 31,913 Da. The N-terminal 56 amino acids (comprising the DNA-binding region) were removed, yielding a protease-resistant 290 residue domain (residues 57-338 plus EG and C-terminal His_6 tag; predicted mass: 31,911 Da). Similar cleavage patterns were observed for Subtilisin and Proteinase K at protease:protein concentration ratios of 1:10,000 and 1:5000, respectively. After identifying the optimum protease/concentration combinations (Table 1) that produced stable cleavage products, directed in situ proteolysis crystallization trials were performed. All three protease/concentration combinations yielded crystals, with one leading to an X-ray structure determined at 2.45 Å resolution (PDB code: 3DBI). The remaining nine combinations were not pursued because either the protein was not fully digested to form a single stable cleavage product or it was completely degraded.

Our strategy in surveying protease/concentration combinations permitted efficient identification of a few optimal combinations for the protein target that subsequently yielded crystals. In this particular case, the presence of the N-terminal region appears to have hindered crystallization of full-length protein. Appropriately chosen conditions for in situ proteolysis per-

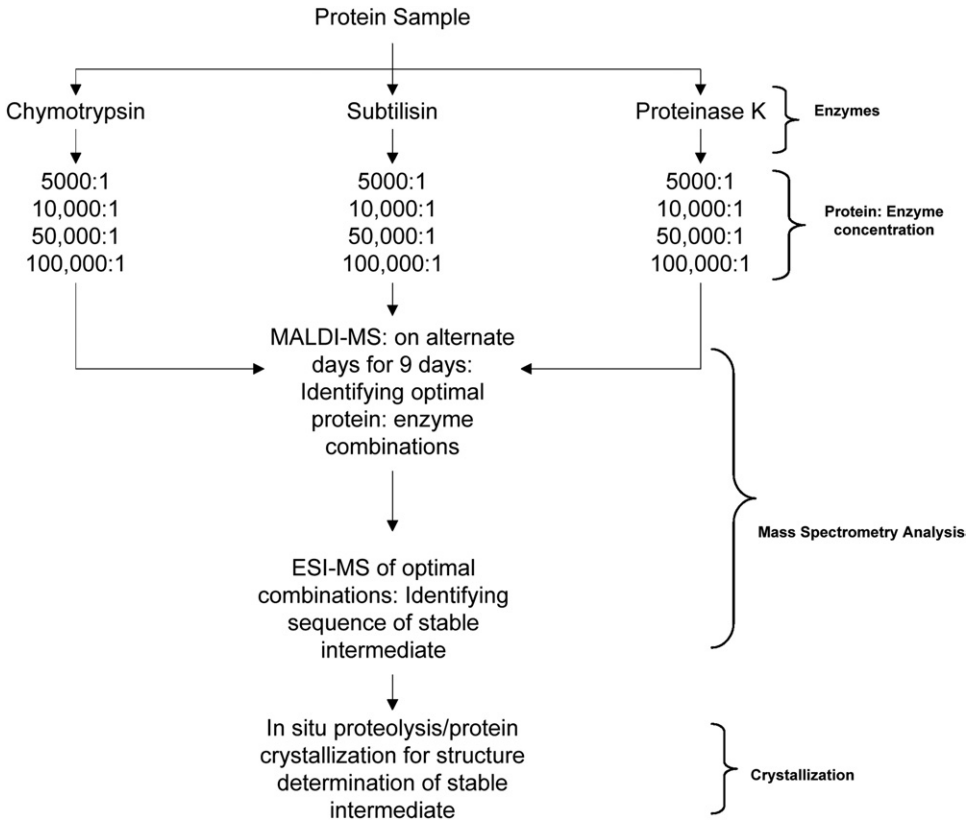

Figure 2. Protocol for mass spectrometry guided in situ proteolysis to obtain crystals for structure determination.

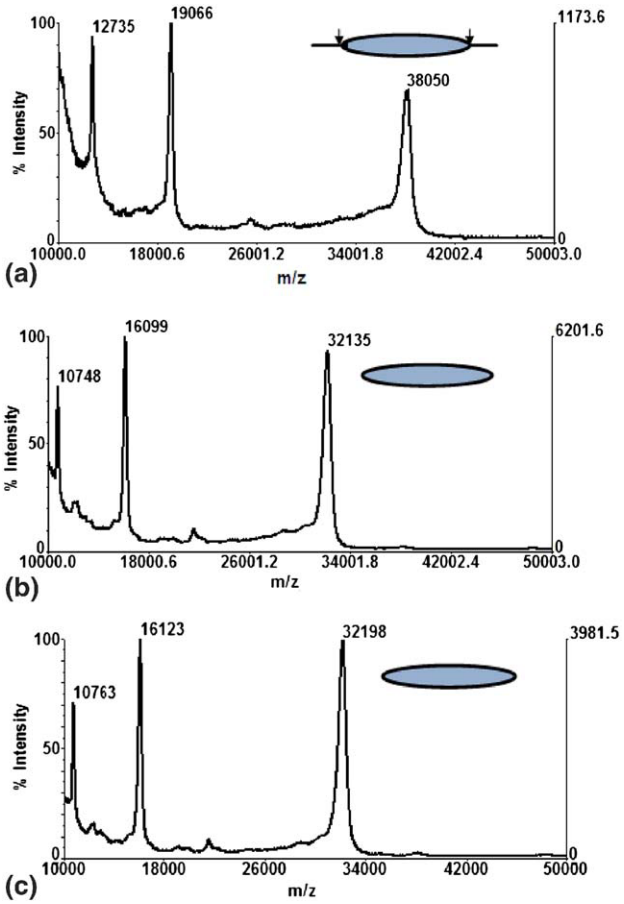

Figure 3. MALDI mass spectra of an *E. coli* transcriptional repressor, Genbank NP_417194, PSI TargetID "NYSGXRC-11030h" (**a**) without addition of any protease; (**b**) after incubating with chymotrypsin (1:5000) for 1 day; (**c**) after incubating with chymotrypsin (1:5000) for 9 d. See Figure 1 for explanation of inset cartoon.

mitted removal of this segment followed by successful nucleation and formation of diffraction quality crystals.

Internal loop digestion can also yield diffraction quality crystals [24, 25]. One of our 10 recalcitrant protein targets succumbed following internal loop cleavage followed by complete digestion of a small C-terminal segment. The predicted mass of this target (Genbank NP_416870, NYSGXRC-11029y, NYSGXRC proteinID 15546) is 24004 Da. Figure 4a shows the intact mass analysis of the full length protein before protease addition. After 1 days of digestion with proteinase K (1:10,000), two peaks were observed at 9.4 and 14.5 kDa (Figure 4b). Thereafter, the polypeptide chain segment corresponding to the peak at 9.4 kDa was degraded, resulting in a single 14.5 kDa peak that resisted further cleavage for 8 additional days (Figure 4c). Similar findings were obtained with Subtilisin at a 1:5000 protease:protein ratio. Using ESI-MS, the stable portion of the protein was identified as the N-terminal domain comprising 135 amino acids; the C-terminal portion (79 amino acids) having been completely removed and then degraded. These two protease conditions were used during subsequent directed in situ proteolysis crystallization trials and both these conditions yielded crystals, one of which resulted in a crystal structure determined at 1.45 Å resolution (PDB code: 3F6C).

In situ proteolysis has previously been shown to be a successful salvage technique for obtaining better quality crystals [4, 18]. When pursuing high-throughput structure determination, however, choosing a single protease and relative enzyme concentration may not be sufficient. Table 1 documents that for a given target only a few of the twelve surveyed protein:protease combinations yielded stable, homogenous (SH) proteolysis products that are suitable for subsequent crystallization trials. The remaining protein:protease combinations need not be examined further. Using the protocol described in this paper, we were able to triage in situ proteolysis crystallization conditions for each target, thereby saving considerable futile efforts in the crystallization laboratory. Moreover, four of our 10 structure determination targets were shown to be unsuitable for in situ proteolysis crystallization. Target NYSGXRC-9269c, an amidohydrolase, (PID: 14771) was completely degraded by each of the three proteases at all protease:

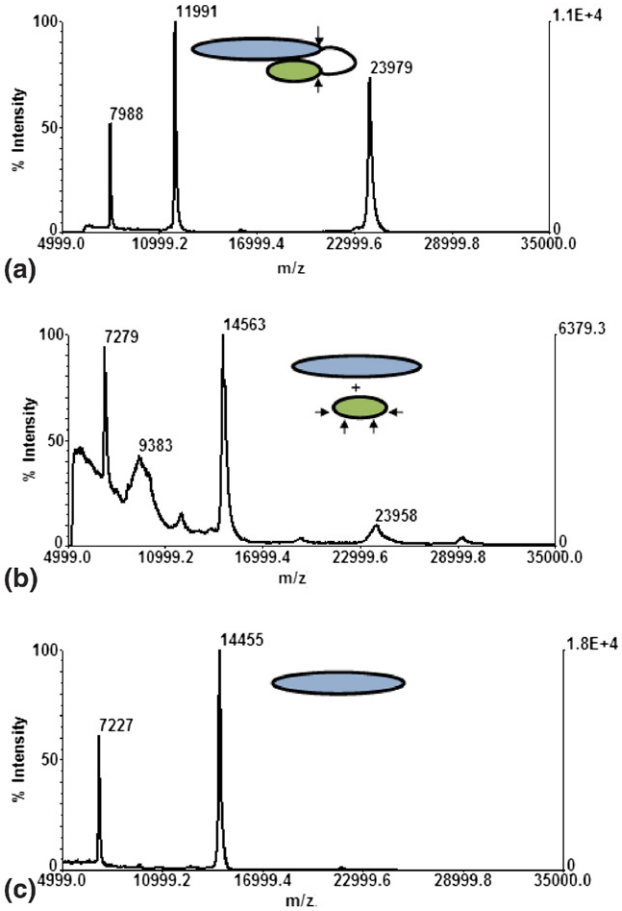

Figure 4. MALDI mass spectra of protein sample (Genbank NP_416870, NYSGXRC-11029y, NYSGXRC proteinID 15546) (**a**) without addition of any protease; (**b**) after incubating with proteinase K (1:10,000) for 1 d; (**c**) after incubating with proteinase K (1:10,000) for 9 d. See Figure 1 for explanation of inset cartoon.

protein ratios used in the study. An additional three targets underwent massive precipitation within 24 h of protease addition and failed to yield interpretable mass spectra.

In summary, application of our strategy yielded two novel protein structures at a total "cost" of 12 sets of crystallization attempts on six apparently promising targets. Brute force in situ proteolysis with the same protease:protein combinations for all 10 targets would have yielded the same two novel structures, at a 10-fold higher total "cost" of 120 sets of crystallization trays.

Conclusions

To support high-throughput X-ray structure determination of proteins, our enzymatic digestion/mass spectrometry strategy permitted identification of optimal combinations of protease and target protein for crystallization *via* proteolysis/MS. Proteins tend to respond differently to protease exposure based on their primary, tertiary and quaternary structures. We have demonstrated that identifying optimal experimental conditions before directed in situ proteolysis crystallization trials can significantly improve the efficiency of this structure determination salvage pathway. Our protocol rapidly identifies experimental conditions most likely to produce a homogenous, stable protein species. This approach not only eliminates need for recloning, expression, and purification (as with LPMS), but also increases the likelihood of identifying conditions that yield limited digests of challenging protein targets. Finally, it can be used to identify proteins with intrinsically disordered regions that are largely or completely digested in the presence of trace amounts of protease, suggesting that alternative methods may be necessary to determine their 3D structure.

Acknowledgments

The authors thank Dr. Devon Thompson for valuable suggestions and the NYSGXRC protein production team. The material is based on research supported by the National Institutes of Health (grant U54 GM074945; PI: S.K.B.). This work was carried out at SGX Pharmaceuticals, Inc. and Eli Lilly and Company, which acquired SGX. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Sciences. Use of the LRL-CAT beam line facilities at Sector 31 of the APS was provided by Eli Lilly and Company, which operates the facility.

References

1. Bierzynski, A. Methods of Peptide Conformation Studies. *Acta Biochim. Polonica* **2001,** *48,* 1091–1099.
2. Cai, X.; Dass, C. Conformational Analysis of Proteins and Peptides. *Curr. Org. Chem.* **2003,** *7,* 1841–1854.
3. Pusey, M. L.; Liu, Z. J.; Tempel, W.; Praissman, J.; Lin, D.; Wang, B. C.; Gavira, J. A.; Ng, J. D. Life in the Fast Lane for Protein Crystallization and X-Ray Crystallography. *Prog. Biophys. Mol. Biol.* **2005,** *88,* 359–386.
4. Wernimont, A.; Edwards, A. In Situ Proteolysis to Generate Crystals for Structure Determination: An Update. *PLoS One* **2009,** *4,* e5094.
5. Grabowski, M.; Joachimiak, A.; Otwinowski, Z.; Minor, W. Structural Genomics: Keeping up with Expanding Knowledge of the Protein Universe. *Curr. Opin. Struct. Biol.* **2007,** *17,* 347–353.
6. O'Toole, N.; Grabowski, M.; Otwinowski, Z.; Minor, W.; Cygler, M. The Structural Genomics Experimental Pipeline: Insights from Global Target Lists. *Proteins* **2004,** *56,* 201–210.
7. Dale, G. E.; Oefner, C.; D'Arcy, A. The Protein as a Variable in Protein Crystallization. *J. Struct. Biol.* **2003,** *142,* 88–97.
8. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization-Principles and Practice. *Mass Spectrom. Rev.* **1990,** *9,* 37–70.
9. Karas, M.; Hillenkamp, F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons. *Anal. Chem.* **1988,** *60,* 2299–2301.
10. Tanaka, K.; Waki, H.; Ido, H.; Akita, S.; Yoshida, T. Protein and Polymer Analyses up to *m/z* 100,000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988,** *2,* 151–153.
11. Geerlof, A.; Brown, J.; Coutard, B.; Egloff, M. P.; Enguita, F. J.; Fogg, M. J.; Gilbert, R. J. C.; Groves, M. R.; Haouz, A.; Nettleship, J. E.; Nordlund, P.; Owens, R. J.; Ruff, M.; Sainsbury, S.; Svergun, D. I.; Wilmanns, M. The Impact of Protein Characterization in Structural Proteomics. *Acta Crystallogr.* **2006,** *D62,* 1125–1136.
12. Cohen, S. L.; Chait, B. T. Mass Spectrometry as a Tool for Protein Crystallography. *Annu. Rev. Biophys. Biomol. Struct.* **2001,** *30,* 67–85.
13. Gao, X.; Bain, K.; Bonanno, J. B.; Buchanan, M.; Henderson, D.; Lorimer, D.; Marsh, C.; Julie, A.; Reynes, J.; Sauder, J. M.; Schwinn, K.; Thai, C.; Burley, S. K. High-Throughput Limited Proteolysis/Mass Spectrometry for Protein Domain Elucidation. *J. Struct. Funct. Genomics* **2005,** *6,* 129–134.
14. Zhang, Z.; Smith, D. L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **1993,** *2,* 522–531.
15. Diederichs, K.; Diez, J.; Greller, G.; Muller, C.; Breed, J.; Schnell, C.; Vonrhein, C.; Boos, W.; Welte, W. Crystal Structure of MalK, the ATPase Subunit of the Trehalose/Maltose ABC Transporter of the Archaeon *Thermococcus litoralis*. *EMBO J.* **2000,** *19,* 5951–5961.
16. Zhang, Z.; Post, C. B.; Smith, D. L. Amide Hydrogen Exchange Determined by Mass Spectrometry: Application to Rabbit Muscle Aldolase. *Biochemistry* **1996,** *35,* 779–791.
17. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L. Jr. Protein Structure Change Studied by Hydrogen–Deuterium Exchange, Functional Labeling, and Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100,* 7057–7062.
18. Dong, A.; Xu, X.; Edwards, A. M. In Situ Proteolysis for Protein Crystallization and Structure Determination. *Nature Methods* **2007,** *4,* 1019–1021.
19. Cadene, M.; Chait, B. A Robust, Detergent-Friendly Method for Mass Spectrometric Analysis of Integral Membrane Proteins. *Anal. Chem.* **2000,** *72,* 5655–5658.
20. Leslie, A. G. W.; Brick, P.; Wonacott, A. J. An Improved Program Package for the Measurement of Oscillation Photographs. *CCP4 News* **1986,** *18,* 33–39.
21. Bailey, S. Collaborative Computing Project Number 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **1994,** *50,* 760–763.
22. Lehninger A. L.; Nelson, D. L.; Cox, M. M. Lehninger Principles of Biochemistry. Worth Publishers Inc.: New York, NY, 2000; p. 60.
23. Sauder, J. M.; Rutter, M. E.; Bain, K.; Rooney, I.; Gheyi, T.; Atwell, S.; Thompson, D. A.; Emtage, S.; Burley, S. K. High Throughput Protein Production and Crystallization at NYSGXRC. *Methods Mol. Biol.* **2008,** *426,* 561–575.
24. Mandel, C. R.; Gebauer, D.; Zhang, H.; Tong, L. A Serendipitous Discovery That In Situ proteolysis is Required for the Crystallization of Yeast CPSF-100 (Ydh1p). *Acta Crystallogr.* **2006,** *F62,* 1041–1045.
25. Xiang, S.; Usunow, G.; Lange, G.; Busch, M.; Tong, L. Crystal Structure of 1-Deoxy-d-xylulose 5-Phosphate Synthase, a Crucial Enzyme for Isoprenoids Biosynthesis. *J. Biol. Chem.* **2006,** *282,* 2676–2682.
26. Sicheri, F.; Moarefi, I.; Kuriyan, J. Crystal Structure of the Src Family Tyrosine Kinase Hck. *Nature* **1997,** *385,* 602–609.