Analytical Biochemistry 488 (2015) 6–8

Contents lists available at ScienceDirect

Analytical Biochemistry

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

Notes & tips

Direct N-terminal sequencing of polypeptides using a thermostable bacterial aminopeptidase and MALDI-TOF mass spectrometry

Nitin Kishor, Purnananda Guptasarma *

Center for Protein Science, Design and Engineering (CPSDE), Department of Biological Sciences, Indian Institute of Science Education & Research (IISER) Mohali, Knowledge City, Sector-81, SAS Nagar, Punjab, 140306, India

ARTICLE INFO

Article history:
Received 3 June 2015
Accepted 9 July 2015
Available online 21 July 2015

Keywords:
Non-specific aminopeptidase
Processive aminopeptidase
Thermostable aminopeptidase
Bacillus subtilis aminopeptidase
Peptide N-terminal sequencing
MALDI-TOF mass spectrometry

ABSTRACT

Mass spectrometry-based amino acid sequencing is currently based almost entirely on collision-induced peptide fragmentation and analyses. Here, we describe a single-stage MS-based technique for amino acid sequencing involving partial, heterogenous digestion of a peptide by a processive, non-specific, heat-loving Bacillus subtilis-derived aminopeptidase (BsuAP), which acts optimally at 70 °C and allows 'single-shot' sequencing to be carried out through simultaneous accumulation, and detection of sub-populations of peptides of progressively reducing length.

© 2015 Elsevier Inc. All rights reserved.

Aminopeptidases are exo-peptidase enzymes. They catalyze the hydrolysis of the peptide bond which joins the N-terminal amino acid of a peptide substrate to the immediate next amino acid along the peptide's sequence.

1. Categories of aminopeptidases

There are two kinds of aminopeptidases. (i) Residue-specific aminopeptidases recognize and remove only specific N-terminal amino acids (e.g., methionine aminopeptidase removes only methionine). Residue-specific aminopeptidases can usually only remove one amino acid residue from the N-terminus of a peptide since, after such removal, the reaction halts if the next amino acid is a different one which is not recognized by the aminopeptidase. (ii) Non-specific aminopeptidases are capable of removing virtually any N-terminal amino acid. Consequently, such aminopeptidases are capable of performing this function not just for peptides with entirely different sequences, different origins, and different N-termini, but also for peptides that have been generated from the same initial population of peptides of identical sequence, through successive removal of different residues. Each removal of a residue generates a peptide with a new N-terminal amino acid. Such aminopeptidases are also called 'processive' aminopeptidases. All aminopeptidases which are non-specific in their activity are potentially processive in nature. Some have been reported to cleave and separate up to 30–35 residues from the N-termini of different substrate polypeptides [1,2].

2. The concept of using an aminopeptidase for peptide sequencing

The ability of a non-specific aminopeptidase to be 'processive' creates an interesting theoretical possibility. Since all molecules in any substrate peptide population of a given amino acid sequence are unlikely to be acted upon simultaneously, or with similar efficiencies, the action of an aminopeptidase could potentially give rise to a heterogenous mixture of peptides of different lengths and sizes, starting from a population of peptides of identical size, sequence and length. This gives rise to the possibility of the use of mass spectrometric methods to distinguish between peptides of differing lengths, followed by a 'bioinformatics-based' inference of the sequence of amino acids – through simple analysis of differences in masses of peptides of progressively reducing size and length. Beginning with a homogenous population of a defined mass constituting a single amino acid sequence, a processive aminopeptidase can potentially create a population of peptides of all possible lengths, differing by the mass of one (or more) of the twenty naturally-occurring amino acids. Such a heterogenous population could be analyzed in a single mass spectrum, e.g., a

* Corresponding author.
E-mail address: guptasarma@iisermohali.ac.in (P. Guptasarma).

http://dx.doi.org/10.1016/j.ab.2015.07.006
0003-2697/© 2015 Elsevier Inc. All rights reserved.

MALDI-TOF mass spectrum, to determine the original peptide's N-terminal amino acid sequence. The possibility has long been recognized, and the concept itself is by no means new.

3. Problems in the development of a feasible method

There are multiple problems in developing a feasible method: (1) In practice it has proved to be difficult to control the aminopeptidase reaction and prevent it from progressing too rapidly, i.e., to keep the enzyme from rapidly progressing through the N-terminal residues of most intermediate species generated during the reaction. This leaves residual peptides of only very short peptide lengths, and peptides of intermediate sizes fail to be detected because they have all been digested and shortened by the enzyme. (2) Another problem is that some naturally-occurring peptides have blocked N-termini, and these prevent facile proteolysis. (3) Sometimes a certain amount of local three-dimensional structure in the peptide/polypeptide could reduce the processivity of the enzyme, and interfere with the progress of the cleavage reaction, causing it to swing from being too rapid to not occurring at all. All of these potential problems have resulted in a situation in which there exists today not a single aminopeptidase-based (mass spectrometry-based) method for sequencing polypeptides, in which a single mass spectrum can be seen to contain peptide species of all conceivable lengths, beginning with the N-terminus.

4. Requirements for developing a feasible method

What is required for the development of a feasible method is an enzyme reagent capable of: (a) working at high temperatures (to allow unfolding of local structures within peptide substrates), (b) 'deblocking' blocked amino termini, and also (c) working slowly-enough, and generating and allowing molecules of all possible peptide lengths to accumulate (owing to differential preferences for different N-terminal amino acids, and therefore different speeds of reaction).

5. Introducing the BsuAP aminopeptidase

In this article, we describe the use of a non-specific, processive, deblocking, *Bacillus subtillis*-derived aminopeptidase enzyme (BsuAP). BsuAP was cloned, expressed and characterized by us previously [3], as a control and template for certain protein engineering experiments involving a comparison with another aminopeptidase from a hypethermophile organism. We discovered that BsuAP has N-terminal de-blocking capability, and that it is not just thermostable but also thermophilic in character and activity. Despite its origins in a mesophile proteome, BsuAP's optimal temperature of activity is 70 °C, [3]. Notably, before our studies, BsuAP was already classified as a potential aminopeptidase, or glucanase, based on bioinformatics analysis, and the structure of the enzyme had also been solved by a structural genomics group which produced and crystallized it without performing further characterization [4]. The structure of BsuAP bears the identifier PDB ID: 1 VHE. We demonstrate that BsuAP is an ideal aminopeptidase for N-terminal sequencing of peptides. Reactions can be conducted at 70 °C, with detection of up to eleven different masses corresponding to eleven possible lengths of a fourteen residues-long polypeptide, beginning with its N-terminus.

6. The sequencing experiment

The peptide substrate used for the experiment reported here was the well-known Glu fibrino peptide (GFP B) which is used as a

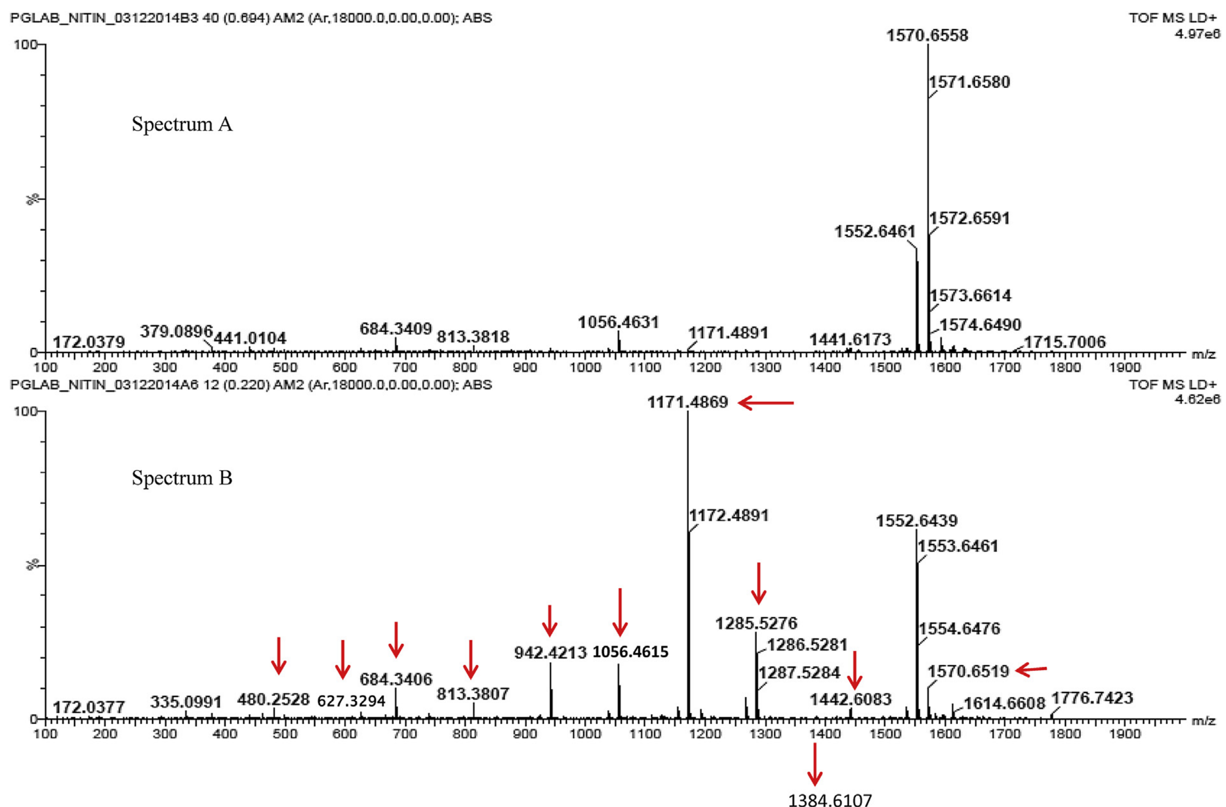

Fig. 1. Spectrum A corresponds to the control reaction lacking BsuAP. Spectrum B corresponds to the reaction of GFP B with BsuAP. The peak intensities of the spectra are 4.98E6 and 4.63E6, respectively. Arrows in Spectrum B denote expected peptide matches for progressively truncated GFP B, beginning with the full-length peptide and labeling peaks for progressively shorter peptides.

Table 1
The sequences and masses of peptides of different length, which were generated by the BsuAP aminopeptidase from the peptide substrate, GFP B.

Peptide sequence	Expected mass	Peptide sequence	Expected mass
EGVNDNEEGFFSAR	1570.677 Da	EGFFSAR	813.39 Da
GVNDNEEGFFSAR	1441.63 Da	GFFSAR	684.35 Da
VNDNEEGFFSAR	1384.61 Da	FFSAR	627.32 Da
NDNEEGFFSAR	1285.54 Da	FSAR	480.25 Da
DNEEGFFSAR	1171.50 Da	SAR	333.19 Da
NEEGFFSAR	1056.47 Da	AR	246.16 Da
EEGFFSAR	942.43 Da	R	175.12 Da

mass calibrating standard in all mass spectrometers. In this case, it was obtained from Waters, USA. GFP B has a mass of 1570.67 Da. The enzyme, BsuAP, was obtained as a recombinant protein expressed in, and purified from, *Escherichia coli* in our own laboratory, as previously described [3]. The reaction mixture of the enzyme and substrate was prepared by mixing 1 nmol of the aminopeptidase to 1.6 μmol of the substrate (GFP B) in water, using a total volume of 100 μL. For the representative experiment, for which results are shown in Fig. 1, the reaction mixture and the control reaction (i.e., the solution lacking aminopeptidase) were separately incubated at 70 °C for 1 h. After incubation, the reaction-mixture and the control were spotted onto a MALDI plate. GFP B was used as a lock mass control. For spotting, 1 μl of sample was spotted with 1 μl CHCA matrix. Samples were analyzed on a dual source (ESI/MALDI) LCMS-Q-TOF system, the Synapt G2S HDMS (from Waters, USA), using the MALDI ionization source and laser (355 nm). The spectra analyzed and presented in Fig. 1 are 'lock mass-corrected' spectra.

7. A satisfactory outcome

The GFP B peptide is 14 residues-long. We explored the assumption that the reaction of the aminopeptidase upon this peptide would not proceed to completion on all substrate molecules, and that a sufficiently slow reaction could produce a mass spectrum containing the masses corresponding to (i) the complete peptide, and (ii) all smaller masses obtained by progressively truncating the peptide from its N-terminus. We found that we could indeed detect each and every expected mass down to 480.25 Da (see Fig. 1), as listed in Table 1. The data presented in Fig. 1 and Table 1, may be viewed together with Supplementary Table 1 which presents data on mass peak intensities in control samples (intact peptide) and aminopeptidase-treated samples of GFP B, establishing that for every peptide that was detected, the intensity and total ion count data were far greater for the treated sample than for the control sample. The data suggests that BsuAP is non-specific, and that it progressively digests up to 11 amino acid residues 'deep' in the GFP B peptide substrate, beginning with the N-terminal glycine, without completely digesting any particular peptide population generated during the reaction. In the mass spectrum of the aminopeptidase-treated sample, therefore, masses are seen to have different relative intensities which could be indicative of differences in specificity and reaction efficiencies. In Fig. 1, differences in masses of successive mass peaks (moving leftwards from the control GFP B mass) correspond accurately to the masses of the amino acids removed from each progressively-truncated peptide. It appears, therefore, that BsuAP could be viably used for N-terminal sequencing of peptides by any research group in possession of only MALDI-TOF MS equipment.

8. Future perspectives

We welcome others to try and use this method, and will be happy to share protein samples and/or the clone of this aminopeptidase, so that eventually the efficacy of this enzyme can be tested for a wide variety of substrate peptides.

Acknowledgments

PG thanks IISER Mohali and the Ministry of Human Resource Development, Govt. of India for Grant No. MHRD-14-0064 (Centre of Excellence in Frontier Areas of Science and Technology) given to IISER Mohali for a Center for Protein Science, Engineering and Design (CPSDE). NK thanks the Universities Grants Commission, Govt. of India for a doctoral research fellowship.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ab.2015.07.006.

References

[1] B. Franzetti, G. Schoehn, J.-F. Hernandez, M. Jaquinod, R.W.H. Ruigrok, G. Zaccai, Tetrahedral aminopeptidase: a novel large protease complex from archaea, EMBO J. 21 (2002) 2132–2138.

[2] M.A. Dura, V. Receveur-Brechot, J.-P. Andrieu, C. Ebel, G. Schoehn, A. Roussel, B. Franzetti, Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon Pyrococcus horikoshii, Biochemistry 44 (2005) 3477–3486.

[3] D. Kapoor, B. Singh, S. Karthikeyan, P. Guptasarma, A functional comparison of the TET aminopeptidases of P. furiosus and B. subtilis with a protein-engineered variant recombining the former's structure with the latter's active site, Enzyme Microb. Technol. 46 (2010) 1–8.

[4] J. Badger, J.M. Sauder, J.M. Adams, S. Antonysamy, K. Bain, M.G. Bergseid, S.G. Buchanan, M.D. Buchanan, Y. Batiyenko, J.A. Christopher, S. Emtage, A. Eroshkina, I. Feil, E.B. Furlong, A.S. Gajiwala, X. Gao, D. He, J. Hendle, A. Huber, K. Hoda, P. Kearins, C. Kissinger, B. Laubert, H.A. Lewis, J. Lin, K. Loomis, D. Lorimer, G. Louie, M. Maletic, C.D. Marsh, I. Miller, J. Molinari, H.J. Muller-Dieckmann, J.M. Newman, B.W. Noland, B. Pagarigan, F. Park, T.S. peat, K.W. post, S. Radojicic, A. Ramos, R. Romero, M.E. Rutter, W.E. Sanderson, K.D. Schwinn, J. Tresser, J. Winhoven, T.A. Wright, L. Wu, J. Xu, T.J. Harris, Structural analyses of a set of proteins resulting from a bacterial genomics project, Proteins 60 (2005) 787–796.