Journal of Pharmaceutical and Biomedical Analysis 55 (2011) 574–582

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

# Quantitative evaluation of protein conformation in pharmaceuticals using cross-linking reactions coupled with LC–MS/MS analysis

Hideto Yamaguchi[a,d,*], Yutaka Hirakura[b], Hiroki Shirai[c], Hisashi Mimura[d], Toshimasa Toyo'oka[a]

[a] Laboratory of Analytical and Bio-Analytical Chemistry, School of Pharmaceutical Sciences and Global COE Program, University of Shizuoka, 52-1 Yada, Suruga-ku, Shizuoka 422-8526, Japan
[b] Analytical Science, Analysis & Pharmacokinetics Research Labs, Drug Discovery Research, Astellas Pharma Inc., 21 Miyukigaoka, Tsukuba, Ibaraki 305-8585, Japan
[c] Applied Genomics, Molecular Medicine Laboratories, Astellas Pharma Inc., 21 Miyukigaoka, Tsukuba, Ibaraki 305-8585, Japan
[d] Pharmaceutical Analysis I, Pharmaceutical Research and Technology Laboratories, Astellas Pharma Inc., 180 Ozumi, Yaizu, Shizuoka 425-0072, Japan

ARTICLE INFO

*Article history:*
Received 9 November 2010
Received in revised form 28 January 2011
Accepted 31 January 2011
Available online 1 March 2011

*Keywords:*
Protein folding/refolding
Protein structure
Stability
Tertiary structure
QC testing

ABSTRACT

The need for a simple and high-throughput method for identifying the tertiary structure of protein pharmaceuticals has increased. In this study, a simple method for mapping the protein fold is proposed for use as a complementary quality test. This method is based on cross-linking a protein using a [bis(sulfosuccinimidyl)suberate ($BS^3$)], followed by peptide mapping by LC–MS. Consensus interferon (CIFN) was used as the model protein. The tryptic map obtained via liquid chromatography tandem mass spectroscopy (LC–MS/MS) and the mass mapping obtained via matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy were used to identify cross-linked peptides. While LC–MS/MS analyses found that $BS^3$ formed cross-links in the loop region of the protein, which was regarded as the biologically active site, sodium dodecyl-sulfate polyacrylamide gel electrophoresis demonstrated that cross-linking occurred within a protein molecule, but not between protein molecules. The occurrence of cross-links at the active site depends greatly on the conformation of the protein, which is determined by the denaturing conditions. Quantitative evaluation of the tertiary structure of CIFN was thus possible by monitoring the amounts of cross-linked peptides generated. Assuming that background information is available at the development stage, this method may be applicable to process development as a complementary test for quality control.



## 1. Introduction

Recent progress in the field of biotechnology has made it possible to generate novel types of biotechnology-derived protein pharmaceuticals, such as transgenic/clone animal-derived biopharmaceutical products, as well as conventional recombinant DNA-derived proteins. Technological progress provides increasing opportunities to develop and market new classes of biopharmaceutical products. Analytical technologies must be developed to appropriately evaluate the properties of these complex protein pharmaceuticals to ensure their quality, efficacy, and safety. In turn, improvement of these analytical technologies is essential for the development of protein pharmaceuticals.

It is well known that formation of the 3D structure is the most important element that allows protein pharmaceuticals to exert specific activity [1]. The pharmaceutical industry, however, suffers from a lack of analytical methods for examining the tertiary protein structure for quality control. For the conformational analyses of proteins, several different methods such as nuclear magnetic resonance (NMR) [2], X-ray structural analysis [3] and those based on thermal [4] and fluorescent analysis have widely been used. Although NMR and X-ray analysis provide detailed information on the tertiary structure of protein, the need for an isotope-labeled sample and/or a high quality single crystal sample make these methods less convenient. These highly sophisticated analytical methods are not suitable for use in lot-to-lot quality control, process development and formulation development due to difficulty of sample preparation and the length of time required for the analysis. Thermal and fluorescent analyses are widely used conventional methods, but they provide limited information on tertiary protein structure. The development of a relatively simple analytical method that provides sufficient structural information is required for the structural evaluation and quality control of protein pharmaceuticals.

The development of the mass spectrometric method has progressed rapidly. A series of analytical methods using mass

* Corresponding author at: Pharmaceutical Analysis I, Pharmaceutical Research and Technology Laboratories, Astellas Pharma Inc., 180 Ozumi, Yaizu, Shizuoka 425-0072, Japan. Tel.: +81 54 627 5263; fax: +81 54 627 9821.
*E-mail address:* hyamaguchi@agensys.com (H. Yamaguchi).

0731-7085/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2011.01.038

spectrometry has been developed to investigate tertiary protein structure. For instance, a new method using differences in ionic mobility is capable of discriminating between protein species of different conformations [5]. Methods to determine the fold, such as mass spectrometry coupled with deuterium exchange reactions [6,7] and peptide mapping using various cross-linking reagents and mass spectrometry (MS3D) [8–11], can also detect dynamic properties in a protein molecule. In the MS3D method, proposed by Young et al. [12], a protein molecule in its native tertiary structure is subjected to cross-linking modification using various reagents, followed by enzymatic digestion, after which the modified amino acid residues are subsequently determined by MS analysis. The information on the distance between the modified amino acid residues is collected based on the spacer length of the cross-linking reagent, and is used to provide constraint in determining structure of the protein. During the NMR analysis performed to determine the tertiary structure of the protein molecule, conformational information, such as the distance and dihedral angle between the hydrogen, carbon, and nitrogen atoms of the protein, is used to determine its tertiary structure with the aid of molecular dynamics and mechanistic calculations. However, the MS3D method, which is conceptually similar to NMR, provides less information on the distance between residues and, for this reason, it is known as "low resolution 3D structure mapping" [13]. It should, however, be noted that MS3D offers the advantages of ease and speed with respect to sample preparation and test time.

In this study, the MS3D-based method was applied to consensus interferon (CIFN) with the goal of determining the utility of the test method as a quality control test for protein pharmaceuticals. The theoretical tertiary structure was constructed *in silico* based on the known structures of interferons. CIFN was stored in various denaturing conditions and incubated with a cross-linking agent, followed by tryptic digestion. The cross-linked peptides generated were analyzed using mass spectrometry.

## 2. Experimental

### 2.1. Materials

Recombinant human consensus interferon (CIFN) solution, composed of 27 mM sodium phosphate and 100 mM sodium chloride (pH 7.0), was produced in-house using *E. Coli* as the host cell. Bis(sulfosuccinimidyl)suberate ($BS^3$) was purchased from Pierce Biotechnology (Rockford, IL, USA). 2-Amino-2-hydroxymethyl-1,3-propanediol was purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Prestain marker protein and Coomassie Brilliant Blue (CBB) solution for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) were purchased from Bio-Rad (Alfred Nobel Drive Hercules, CA, USA). The gel for SDS–PAGE, PAG mini "DAIICHI" 15/25, was purchased from DAIICHI Chemical (Tokyo, Japan). Sequencing grade modified Trypsin was purchased from Promega Corporation (Woods Hollow Road Madison, WI, USA). All the other reagents were purchased from Kanto Chemical (Tokyo, Japan).

### 2.2. Preparation of denatured CIFN solutions

The solutions containing CIFN with altered high-order structure were prepared by incubating under various denaturing conditions. To obtain heat-denatured protein solutions, CIFN was incubated for 24 h at different temperatures (37 °C, 50 °C, and 70 °C). To obtain urea denatured solution, CIFN was diluted with an equal volume of 8 M urea and further incubated for 3 h at either 60 °C or 80 °C. These solutions were subjected to dialysis (MW cutoff: 10,000 Da, Pierce, Rockfold, IL, USA) against phosphate buffered saline (PBS) before further analysis. To confirm the degree of protein denaturation, the samples prepared by incubating at 80 °C for 3 h in the presence of 4 M urea (the most hostile conditions in this study) was subjected to the anti-virus cell bioassay.

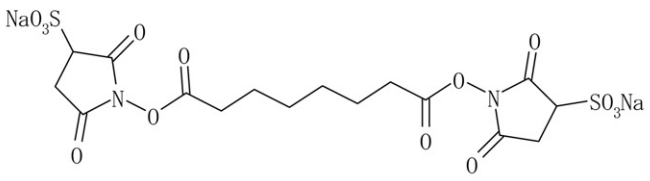
### 2.3. Cross-linking reactions and modification evaluation

Bis(sulfosuccinimidyl)suberate ($BS^3$), used without further purification (Fig. 1a), was dissolved in PBS at a concentration of 20 mM. CIFN solution was diluted with $BS^3$/PBS solution to a protein concentration of 5 μM at CFIN/$BS^3$ ratios (mol/mol) of 1/27, 1/54, and 1/108. These sample solutions were incubated for 24 h at ambient temperature to allow cross-linking reactions to proceed. The reaction was terminated by adding 1/100 volume of 1 M Tris solution (pH 7.4). To monitor the reaction, the sample solution was analyzed using SDS–PAGE in reduced and non-reduced conditions, and liquid chromatography combined with mass spectroscopy (LC–MS). For SDS–PAGE, the sample solution was diluted with a loading buffer with or without 2-mercaptoethanol, and the diluted samples were subsequently transferred onto a 10–20% gradient SDS–PAGE gel. Protein bands were visualized via CBB staining. LC–MS was performed using triple quadrupole mass spectroscopy (API4000, Applied Biosystems/MDS Sciex, Toronto, Canada) equipped with a liquid chromatography system (HP1100, Hewlett Packard, Wilmington, DE, USA). For the MS analysis, samples were desalted using a size-exclusion column (Polyhydroxyethyl A, PolyLC, MD, USA) with a mobile phase (composed of 0.1% formic acid in 50% methanol, at a flow rate 0.5 mL/min). The column temperature and injection volume were set at 25 °C and 100 μL, respectively. The sample was ionized in electrospray ionization (ESI) positive mode. The spray voltage, turbo ion spray temperature, and analyzed MS range were 5.0 kV, 600 °C, and 500–3000 $m/z$, respectively.

### 2.4. Identification of cross-linking site

CIFN and $BS^3$-modified CIFN (CIFN/$BS^3$ molar ratio, 1/54) sample solutions prepared as described in the previous section were dialyzed against PBS (MW cutoff: 10,000 Da, Pierce, Rockfold, IL, USA) and dried under vacuum. The samples were diluted with Tris buffer solution containing guanidinium chloride (GuHCl) at a protein concentration of 1.0 mg/mL. After incubation for 15 min at 37 °C, Dithiothreitol (DTT) solution was added, and the solution was further incubated for 15 min at 37 °C to reduce disulfide bond interactions. Subsequently, monoiodoacetic acid solution was added, which allowed for Cys residue alkylation in the absence of light. The alkylation reaction was terminated by adding an excess volume of 2-mercaptoehanol. After dialysis (MW cutoff, 10,000 Da) against 0.1 M ammonium bicarbonate containing 2 M urea, the samples were digested by adding trypsin at a final trypsin-to-protein ratio of 1:25 (weight/weight) and incubating at 37 °C for 24 h. The digestion reaction was terminated by adding an equal volume of 2% trifluoroacetic acid (TFA) solution to obtain a peptide sample solution.

The peptide sample solution was analyzed on an Agilent 1100 LC system equipped with a Vydac (Hesperia, CA) C18 column (2.1 mm × 250 mm). The elution was performed using a solvent B gradient of 3–60% from 10 to 100 min and 60–90% from 100 to 120 min. Solvent A consisted of 0.1% TFA (Pierce Biotechnology, Rockford, IL, USA) in $H_2O$, and solvent B consisted of 0.1% TFA in 90% acetonitrile (Kanto, Tokyo, Japan) and the flow rate was 0.2 mL/min. The column temperature and injection volume were set at 40 °C and 40 μL, respectively. Fragmented peptides detected in the peaks on the chromatograms were subjected to LC–MS/MS analysis. Q1 scan analysis was conducted using API4000 triple quadrupole mass spectroscopy with a mass range of 400–2000 (spray voltage, 5.0 kV; turbo ion spray, 600 °C). MS/MS analysis was then performed using




(a)

(b)

MCDLPQTHSLGNRRALILLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQ
KAQAISVLHEMIQQTFNLFSTKDSSAAWDESLLEKFYTELYQQLNDLEACVI
QEVGVEETPLMNVDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFS
LSTNLQERLRRKE

**Fig. 1.** The chemical structures are shown. (a) Bis(sulfosuccinimidyl)suberate ($BS^3$), (b) the primary structure of CIFN and sites that could be modified by BS3. The underlined amino acids are the possible modification sites.

the independent data acquisition (IDA) method where the predominant molecular ion in each peak was transferred to MS/MS analysis automatically. The collision energy upon MS/MS analysis was 30 V, 40 V, and 50 V. MALDI-TOF-MS analysis was also performed using ultraflex TOF/TOF (Bruker Daltonics, Billerica, MA, USA). For sample preparation, the peptide sample solution was desalted using a prespotted anchorchip (Bruker Daltonics, Billerica, MA, USA) under predetermined conditions (sample matrix: α-cyano-4-hydroxycinnamic acid (α-CHCA); acceleration voltage: 25 kV; measurement mode: positive reflector mode). $BS^3$-modified peptides were thus identified by comparing tryptic maps of CIFN and $BS^3$-modified CIFN samples obtained using LC–MS with peptide mass fingerprinting with MALDI-TOF-MS. Automated Spectrum Assignment Program (ASAP ver 1.9) [14] developed at the University of California at San Francisco was also employed to identify the cross-linked peptides.

### 2.5. CD analysis

In order to examine whether conformational changes occur during $BS^3$-modification, circular dicroism (CD) analysis was conducted using a spectropolarimeter (model J-720W, Jasco International, Tokyo, Japan) to detect the conformations of CIFN and $BS^3$-modified CIFN. For $BS^3$-modified CIFN, the sample solutions were prepared as described in the section on cross-linking reactions and modification evaluation and subject to dialysis against PBS (MW cutoff: 10,000 Da, Pierce, Rockfold, IL, USA).

### 2.6. Cross-linked peptide quantification

For quantitative evaluation of the conformational changes of denatured CIFN proteins, intact and denatured CIFN protein solutions were prepared according to the section on Preparation of denatured CIFN solution, and subjected to BS3-modification and tryptic digestion as described in the section on Identification of cross-linking site. The peptide samples obtained were separated on an Agilent 1100 LC system equipped with a reverse-phase C18 column (Cadenza, Intact, Kyoto, Japan) (2.1 mm × 100 mm). The column was eluted at 0.2 mL/min using a solvent B gradient of 3–60% from 4 to 40 min and 60–90% from 40 to 100 min. Solvent A consisted of 0.1% TFA in $H_2O$, while solvent B consisted of 0.1% TFA in 90% acetonitrile. The column temperature and injection volume were set at 40 °C and 10 μL, respectively. After separation on the HPLC, $BS^3$-modified peptide samples containing high-order structural information were quantified using the multi-reaction monitoring (MRM) method using API4000 triple quadrupole mass spectroscopy. The samples were ionized via electrospray ionization (ESI) positive mode under predetermined conditions (turbo-gas temperature: 600 °C; ion spray needle voltage nebulizer: 5 kV; gas time: 50 ms; curtain gas time: 40 ms; dwell time for MRM: 170 ms). The combination of precursor and product ions obtained for each $BS^3$-modified peptide was determined using the IDA method described in the section on identification of cross-linking site. To avoid signal deviations resulting from experimental variations, a peptide known to be $BS^3$-modification resistant was used as an internal peptide (IS peptide). Because the precursor ion may contain +2 or +3 charges, the mass number of the product ion was sometimes larger than that of the precursor ion.

### 2.7. Homology modeling of CIFN

A three-dimensional model of CIFN was constructed based on the structure of interferon α-2a (PDB Code: 1ITF) [15]. Molecular operating environment (MOE) graphics software was used to align the sequences of CIFN and interferon α-2a, and to build ten models of CIFN. Three candidate models were chosen from the ten based on the points of low energy and stable conformations on the main and side chains. The most appropriate model was then determined using JOY software [16,17].

## 3. Results and discussion

### 3.1. BS3-modification of CIFN

$BS^3$, a protein modification reagent (Fig. 1a), is known to react with primary amines, such as the nitrogen atoms of the N-terminal amino acid residues and the side chains of lysine residues. As shown in Fig. 1b, CIFN has 10 theoretical reaction sites, including 9 lysine residues within the molecule. Fig. 2 demonstrates the MS spectra of CIFN and $BS^3$-modified CIFN (CIFN/$BS^3$ molar ratios, 1/27 and 1/54). The signal intensities of the MS spectra of the $BS^3$-modified CIFN drastically decreased with the increasing $BS^3$ concentrations. This is attributed to the decrease in the net positive charge on the protein as result of the reaction of $BS^3$ with the positively charged Lysine side chains. It should also be noted that molecular mass changes in CIFN due to $BS^3$-modification were evidenced by the rightward shift in mass numbers. In order to examine whether the secondary structure of CIFN changes during $BS^3$-modification, the proteins were subjected to circular dichroism (CD) analysis (protein/$BS^3$ molar ratio, 1/27). As shown in Fig. 3, no significant differences were observed for CIFN or $BS^3$-modified CIFN, which indicates that the secondary structure of CIFN remained unchanged during the $BS^3$-modification process.

SDS–PAGE analyses showed that the profiles of the samples analyzed in reduced and non-reduced conditions differed significantly (Fig. 4). The electrophoretic profile indicates that the molecular

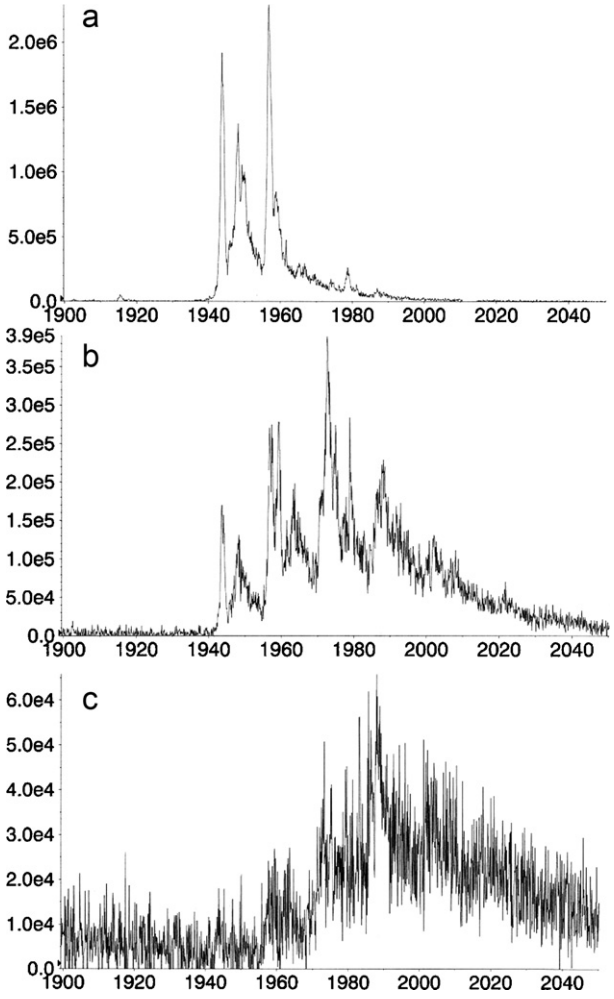



**Fig. 2.** The ESI-MS spectra of CIFN are shown. Each spectrum is the +10 charge state portion of a multi-charge molecular ion spectrum. (a) CIFN, (b) BS3-modified CIFN (protein/BS3 molar ratio, 1/27). (c) BS3-modified CIFN (protein/BS3 molar ratio, 1/54).

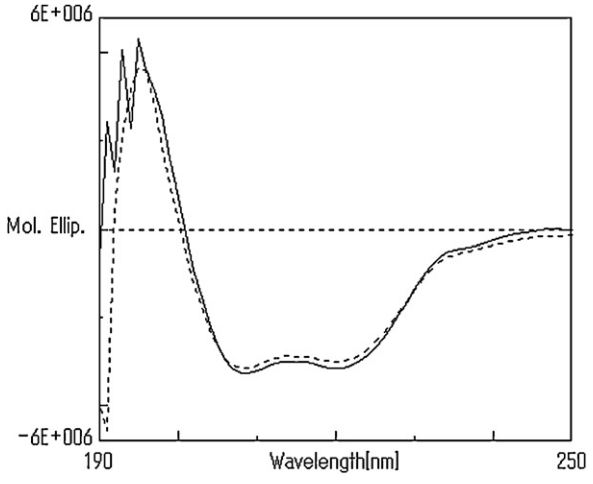



**Fig. 3.** Far-UV circular dichroism spectra of CIFN are shown. Solid line: CIFN; dotted line: BS3-modified CIFN (protein/BS3 molar ratio, 1/27).

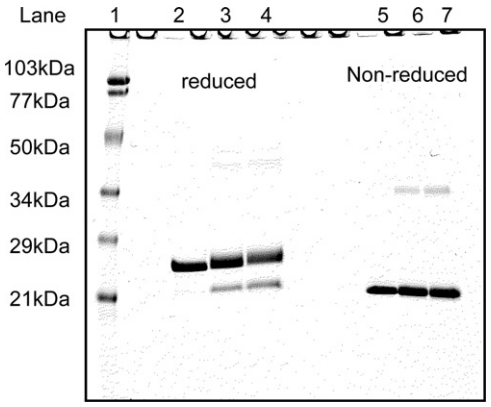



**Fig. 4.** SDS–PAGE profiles of CIFN are shown. Lane 1: marker protein; Lanes 2 and 5: CIFN; Lanes 3 and 6: BS3-modified CIFN (protein/BS3 molar ratio, 1/27); Lanes 4 and 7: BS3-modified CIFN (protein/BS3 molar ratio, 1/54).

weights of the major bands in the non-reduced condition appear to be less than those observed in the reduced condition. This observation may be accounted for by the fact that, in the non-reduced condition, CIFN formed a compact globular structure due to the two internal disulfide bonds, whereas the protein exhibited a relatively disordered structure in the reduced condition. In the non-reduced condition, a band with significantly higher molecular weights was observed, which suggests that $BS^3$ formed inter-molecular cross-links to form a dimer species during the incubation process. While the dimer species should also exist in the reduced condition, broader bands were observed because the reduction of disulfide bonds causes the protein molecules to pack differently, resulting in varying mobility. It should be noted that, in the reduced condition, some molecular species showed the same molecular mobility as the major components in the non-reduced condition. It is highly probable that the $BS^3$-modified CIFN produced in the incubation process carry intra-molecular cross-links, which assist in maintaining the original protein folding conformation. This effect is similar to that provided by an internal disulfide bond, and results in higher mobility. The incubation conditions for $BS^3$ were chosen so that the concentration of protein with intra-molecular cross-links exceeded that of the dimer species.

### 3.2. Determination of the cross-linking sites

ASAP software was used to identify the sequence of all peaks observed in the tryptic map obtained via LC–MS for $BS^3$-modified CIFN. This was followed by MS/MS analysis based on the IDA method and the MS mapping obtained via MALDI-TOF-MS analysis. The average mass determined during tryptic map data analysis was adopted, and the sequence assignment for MS was set at less than 2 Da. The monoisotopic mass determined from the data obtained via MALDI-TOF-MS was adopted, and the sequence assignment for MS was set to be less than 100 ppm. These analyses uncovered five inter-residue cross-linked peptides and two single residue modification sites (Table 1). The results obtained from LC–MS analysis were consistent with those from the MALDI-TOF-MS analysis. MS mapping using data from the MALDI-TOF-MS analyses are shown in Fig. 5 and the locations of the cross-links in the protein are illustrated in Fig. 6. It should be noted that when conducting MS3D, the determination of cross-linked peptides is regarded as the most important and difficult step. Other methods reprinted in the literature include using deuterium- or $^{18}O$-labeled cross-linking reagents [10], or high performance instruments such as Fourier transform ion-cyclotron resonance mass spectrometry (FT-ICR MS) [13,18,19].

**Table 1**
$BS^3$-cross-linked tryptic-digested peptides from consensus interferon as identified by ASAP ver 1.9.

| ESI-MS data | | | MALDI-TOFMS (M+H) | | Assigned sequence | | Residue | Remarks |
|---|---|---|---|---|---|---|---|---|
| Obs. *m*/*z* | Calc. *m*/*z* | Charge | Obs. *m*/*z* | Calc. *m*/*z* | No | Sequence | | |
| 844.0 | 843.3 | 2 | 1685.8 | 1684.8 | −1 to 13 | $BS^3$**M**CDLPQTGTHSLGNR | $M_{-1}$ | |
| 701.0 | 701.5 | 3 | 2102.0 | 2101.1 | 22–33<br>122–126 | ISPFSCL**K**DR<br>**K**YFQR | $K_{31}$–$K_{122}$ | CP#1 |
| 690.4 | 690.2 | 2 | 1379.7 | 1378.7 | 23–33 | ISFPSCL**K**($BS^3$)DR | $K_{31}$ | |
| 920.3 | 919.4 | 3 | N.D. | N.D. | 24–33<br>135–145 | ISPFSCL**K**DR<br>**K**YSPCAWEVVR | $K_{31}$–$K_{135}$ | CP#2 |
| 824.3 | 823.9 | 3 | 2469.3 | 2468.4 | 24–33<br>127–135 | ISPFSCL**K**DR<br>ITLYLTE**KK** | $K_{31}$–$K_{134}$ | CP#3 |
| 833.3 | 832.6 | 3 | 2495.2 | 2494.2 | 127–145 | ITLYLTE**KK**SPCAWEVVR | $K_{134}$–$K_{135}$ | CP#4 |
| 664.0 | 663.4 | 3 | N.D. | N.D. | 122–126<br>127–135 | **K**YFQR<br>ITLYLTE**KK** | $K_{122}$–$K_{134}$ | CP#5 |

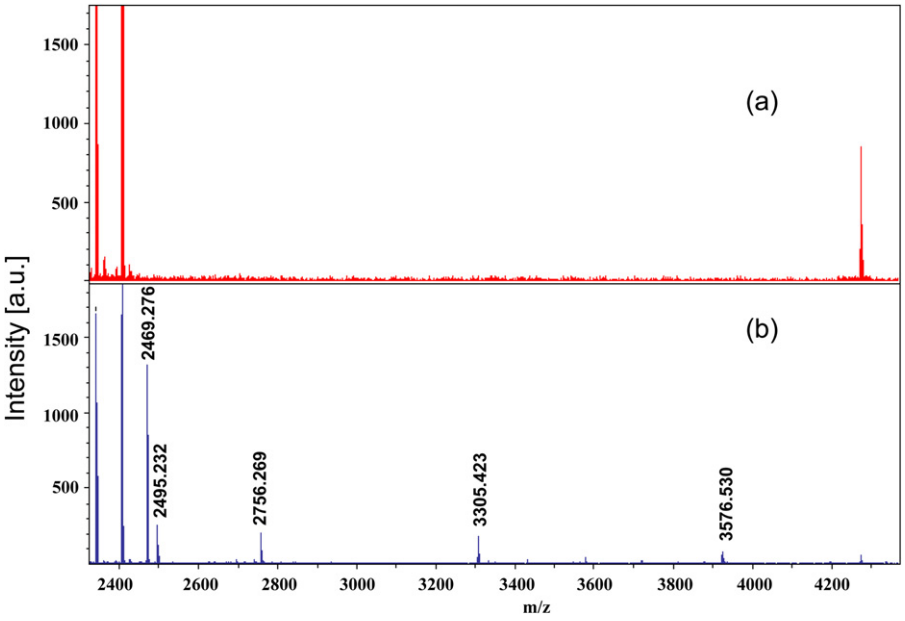



**Fig. 5.** MALDI-TOFMS spectra of tryptic-digested CIFN. (a) CIFN, (b) BS3-modified CIFN (protein/BS3 molar ratio, 1/27). Note that only the fragments of higher mass numbers are illustrated in the mass spectrum of BS3-modified CIFN.

### *3.3. Relationship between the protein conformation and cross-linking sites*

The 3D structures of several subtypes of interferon-α have been determined using NMR and X-ray crystallography; these subtypes are available at Protein Data Bank (PDB). Since the different subtypes of interferon-α belong to the same family, their structures are similar [15]. In the early development stage, it is reasonable to postulate that CIFN has the same conformation as the subtypes of interferon-α. The structure reported for interferon-α-2b (PDB 1ITF) was therefore adopted to construct the 3D structure of CIFN *in silico*, using a typical homology modeling technique [17]. The graphic mapping of those results on the structure produced *in silico* is shown in Fig. 7.

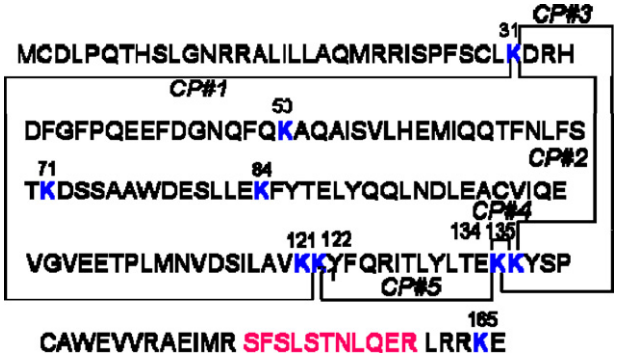



**Fig. 6.** The cross-linking sites and the primary structure of the internal standard (IS) peptide are shown. he IS peptide is derived from the C-terminal region of CIFN. CP#1–CP#5: cross-linked peptides; Red: internal standard (IS) peptide for quantification. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The inter-residual cross-links identified by ASAP (Table 1) were assigned to the postulated three-dimensional structure of the protein (Fig. 7). Although it may be possible that some of the cross-links were derived from two different molecules in a dimer, it appeared that the major part of the inter-residual cross-links originated from a single molecule (Fig. 4). The inter-residue distance for the cross-links between the K31 and K134 residues on the calculated structure was 13.3 Å, which is longer than the $BS^3$ spacer length (11.4 Å). K31 is located on the flexible loop structure, while K134 is located on the flexible position at the terminus of the helix structure. The structural conformation is most likely flexible enough for these residues to form an intra-molecular cross-link. The K134 and K135 residues are adjacent to each other with an inter-residual distance of 8.5 Å. This short distance may facilitate the formation of a cross-link within the molecule. It is also reasonable to postu-

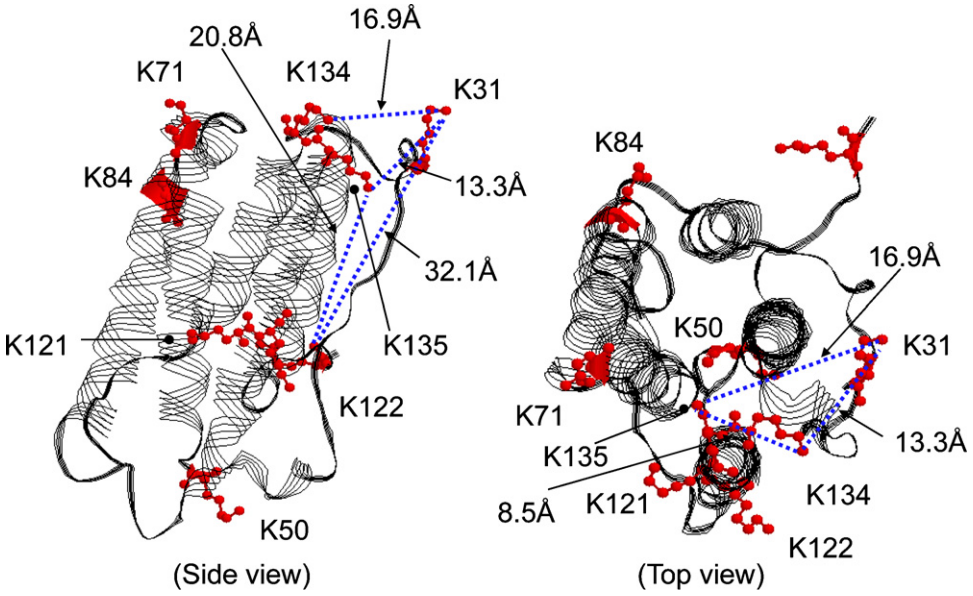



**Fig. 7.** The theoretical 3D structure of CIFN and the assigned cross-links are shown.

late that a cross-link would be formed between K31 and K135 in a single molecule, despite the relatively long inter-residue distance (16.9 Å), which is slightly longer than the $BS^3$ spacer length. The formation of the cross-link can be attributed to the high mobility of K31, as was predicted for the K31 and K134 interactions.

It is unlikely that $BS^3$ will form an intra-molecular crosslink between the K31 and K122 residues because the distance between the residues on the calculated structure is 32.1 Å, which is three times longer than the $BS^3$ spacer length. K122 is located on the central region of the rigid helix, which is where the mobility of K122 would be extremely limited. In the Zn atom-mediated dimer model for Interferon α-2b [20], the distance between the K31 and K122 residues of the different molecules appeared to be significantly shorter than the predicted intra-molecular distance. SDS–PAGE analyses performed on $BS^3$-modified CIFN in the non-reduced condition showed that dimer species did occur (Fig. 4). These findings suggest that the cross-links observed between K31 and K122 residues may be formed inter-molecularly between the different molecules in a dimer species, but not intra-molecularly within a single molecule. Likewise, it is not likely that $BS^3$ forms an intra-molecular cross-link between K122 and K134 because the distance between the residues on the calculated structure is relatively long (20.8 Å). In fact, it is approximately twice as long as the $BS^3$ spacer length. Structural flexibility is also limited for K122 and K134 because both residues are located on the same helix; actually, the cross-link is inter-molecular, not intra-molecular.

It should be noted that, although it was assumed that the low reactivity was due to steric hindrance, which results in limited accessibility by $BS^3$, the mapping results do not support this hypothesis. It is possible that reactivity with $BS^3$ depends on the pKa values of the nitrogen atoms as well as the presence of steric hindrance. Slight differences in the micro-environment around each residue may affect the actual pKa value and lead to reactivity modifications.

Although inter and intra-molecular cross link peptide could be discriminated using tertiary structural modeling of protein, cross-linked monomers should be isolated by size-exclusion chromatography or SDS–PAGE, and the assignment of cross-links should be performed on the isolated species if the method will be applied to proper quality testing.

### 3.4. Quantitative evaluation of cross-linked peptides

Tryptic maps of CIFN and $BS^3$-modified CIFN are shown in Fig. 8 as total ion current chromatograms (TICC). Trypsin cleaves peptide bonds on the C-terminal of arginine and lysine residues. Peaks observed in the TICC of CIFN were identified based on the molecular mass and MS/MS patterns of individual peaks. Intact and denatured CIFN samples that had been stored under various conditions were incubated with $BS^3$ at a protein to $BS^3$ ratio of 1:54 (mol/mol) for 24 h and subsequently processed for tryptic digestion. The cross-linked peptides and an internal standard (IS) peptide resistant to $BS^3$-modification (Fig. 6 and Table 1) were quantified using the MRM method described in the section on cross-linked peptide quantification. The MRM chromatograms of intact CIFN exhibited good specificity for each of the $BS^3$-modified peptides (Fig. 9a). In the MRM chromatograms for denatured CIFN, those peaks derived from the $BS^3$-modified peptides disappeared, whereas the peak for the IS peptide was comparable to that for intact CIFN. Thus, this method could detect the tertiary structure of the protein (Fig. 9b). The peak areas for the five different cross-linked peptides, obtained by incubating intact and denatured CIFN with $BS^3$ and trypsin digestion, were normalized against the peak area for the IS peptide for quantitative analysis. The peak area

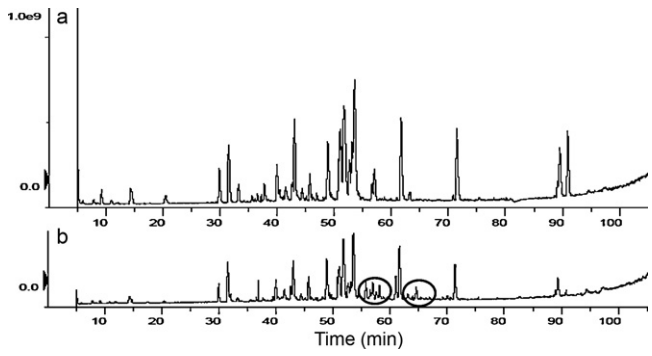



**Fig. 8.** Total ion current chromatograms (tryptic maps) of CIFN are shown. (a) CIFN, (b) BS3-modified CIFN (protein/BS3 molar ratio, 1/27). New peaks derived from BS3-modification in the two regions are indicated by the black circles.

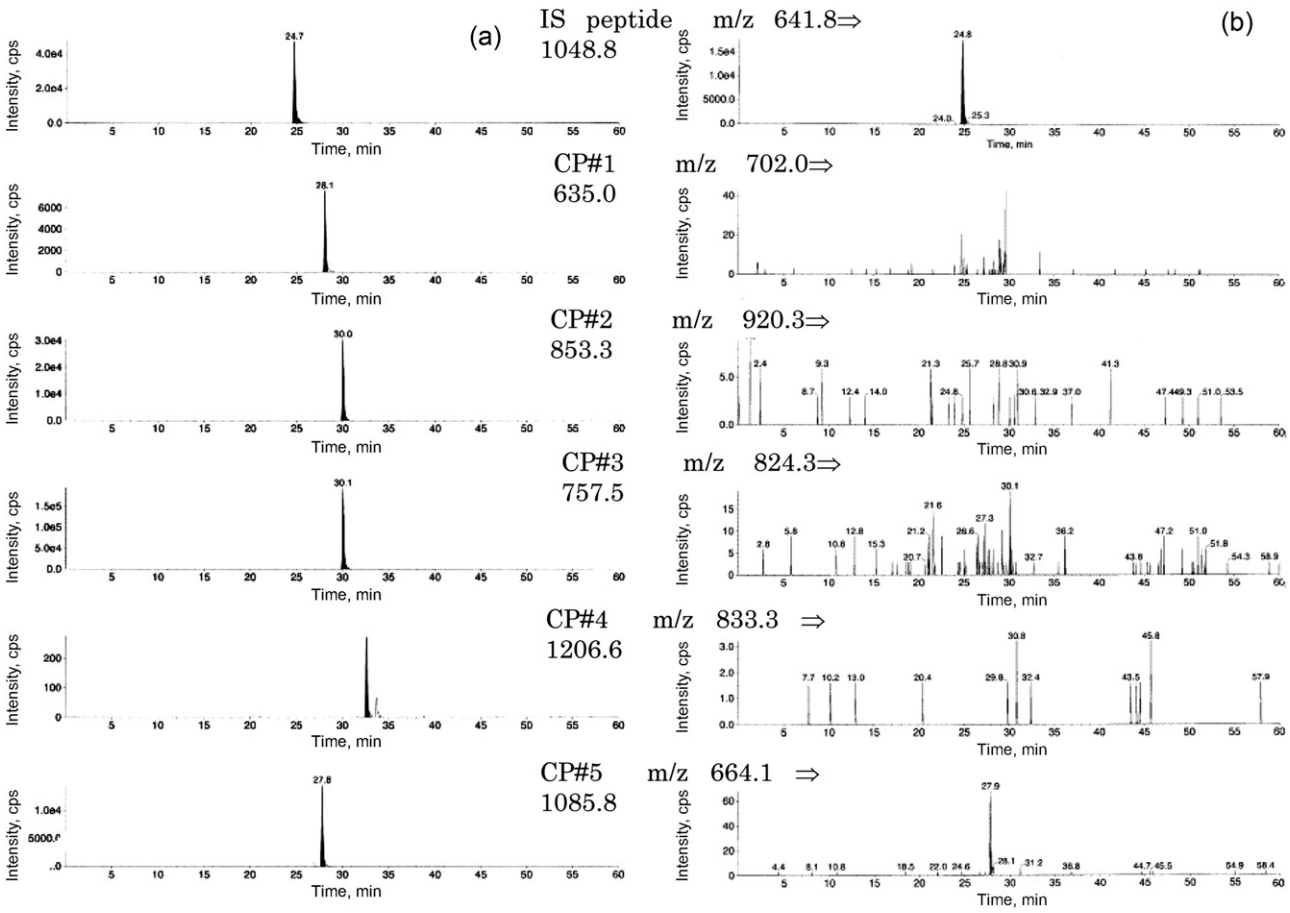



**Fig. 9.** MRM chromatograms of the IS peptide and BS3-modified cross-linked peptides obtained from CIFN are shown. The vertical axes indicate the intensities of product ions. (a) Intact CIFN, (b) denatured CIFN incubated for 3 h at 80 °C with 4 M urea.

ratios of each cross-linked peptide obtained for CIFN that denatured under different conditions were again normalized against the peak area ratio obtained for intact CIFN to produce the relative ratios (Table 2). It should be noted that all cross-linked peptides had lower relative ratios after denaturation. This tendency appears to depend on the denaturing conditions. Three samples, including K31–K135 (CP#2), K31–K134 (CP#3), and K122–K134 (CP#5) cross-linked peptides, showed gradual decreases in their relative ratios when subjected to relatively mild denaturing conditions (i.e. incubation at 37 °C and 50 °C for 24 h). Interestingly, the relative ratio of K31–K122 (CP#1) obtained after incubation at 37 °C increased markedly compared to the other peptides, while at 50 °C, its relative ratio was similar to the other peptides. It is highly likely that the striking increase observed at 37 °C is due to facilitation of inter-molecular cross-link formation in the dimer species at a temperature where the proteins remained intact. Under incubation conditions of 70 °C for 24 h, which were regarded as more hostile conditions, the intra-molecular cross-linked peptides (K31–K135 (CP#2), K31–K134 (CP#3), and K134–K135 (CP#4)) had relative ratios equal to 7–15% of the intact sample, while those of inter-molecular cross-linked peptides (K31–K122 (CP#1) and K122–K134 (CP#5)) were 44% and 31%, respectively, of the intact sample. Under incubation conditions of 60 °C for 3 h in the presence of 4 M urea, which were regarded as more hostile conditions, the intra-molecular cross-linked peptides (K31–K135 (CP#2), K31–K134 (CP#3), and K134–K135 (CP#4)) had relative ratios equal to 6–9% of the intact sample, while those of inter-molecular cross-linked peptides (K31–K122 (CP#1) and K122–K134 (CP#5)) were 37% and 17%, respectively, of the intact sample. The samples prepared by incubating at 80 °C for 3 h in the presence of 4 M urea (the most hostile conditions in this study), which exerted no antivirus activity (data not shown), showed no peaks derived from the cross-linked peptides except for K122–K134 (CP#5). It should be noted that some denaturing conditions dedicated for this study potentially causes chemical modification on the protein, especially incubation in the presence of 4 M urea at 80 °C could exert carbamoylation on α amino groups. Meanwhile, these kinds of chemical modification can be simply detected and quantified by a general peptide mapping analysis coupled with mass spectrometry. So the potential loss of cross-linking reactivity by chemical modifi-

**Table 2**
Relative ratios of the BS3-modified cross-linked peptides obtained from denatured CIFN against intact CIFN (Mean ± SD%).

| Cross linked peptides | 37 °C for 24 h | 50 °C for 24 h | 70 °C for 24 h | 60 °C for 3 h with 4 M urea | 80 °C for 3 h with 4 M urea |
|---|---|---|---|---|---|
| CP#1 | 239 ± 1.4 | 112 ± 4.3 | 44 ± 2.7 | 37 ± 4.9 | N.D. |
| CP#2 | 82 ± 3.5 | 91 ± 3.5 | 13 ± 6.6 | 9 ± 2.7 | N.D. |
| CP#3 | 83 ± 1.7 | 83 ± 0.7 | 15 ± 1.1 | 7 ± 0.8 | N.D. |
| CP#4 | 98 ± 5.5 | 137 ± 5.5 | 7 ± 4.1 | 6 ± 0.1 | N.D. |
| CP#5 | 63 ± 0.7 | 57 ± 0.4 | 31 ± 9.9 | 17 ± 0.1 | 1 ± 0.3 |

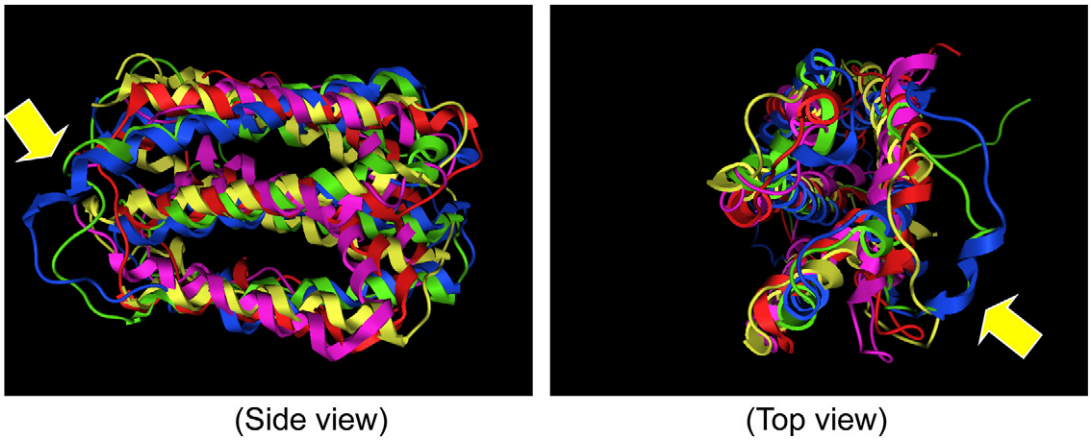



**Fig. 10.** The 3D structures of interferon family (interferon/interleukin-10 family) are shown. A yellow arrow indicates the loop region, which contains K31 of CIFN. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

cation like carbamylation of lys residues can be distinguished from loss of cross-linking reactivity derived from protein unfolding.

### 3.5. Identification tests for 3D structure of proteins

The intra-molecular cross-links detected in this study were assigned to logical locations on the theoretical model structure. It is, therefore, possible to use this method to determine a low resolution tertiary structure of protein pharmaceuticals using cross-links. Quantitative evaluation of the signal intensities of cross-linked peptides may be used as an identification test for tertiary protein structure as well as a simple evaluation test for protein denaturation. However, the following requirements must be satisfied first: first, a three-dimensional structure is available so that cross-links may be assigned to the linking sites; second, incubation with a spacer does not result in denaturation of the protein; and, third, inter-residue cross-links must be formed at sites relevant to biological activity. The method described in this study may provide a potential advantage as a novel analytical method for both qualitative and quantitative assessment of the tertiary structure of protein pharmaceuticals.

Where Cys29 and Cys139 are close enough to interact efficiently, a disulfide bond between the two contributes to the maintenance of the intact conformation of CIFN. In this study, three intra-molecular cross-linked peptides (K31–K134, K31–K135, and K134–135) were formed in the loop region (Fig. 7). Indeed, the structural integrity of the high-mobility molecular species, observed during SDS–PAGE analyses in the reduced condition (Fig. 4), may be maintained in the presence of those intra-molecular cross-links. It is therefore highly likely that the occurrence of two of the intra-molecular cross-linked peptides (K31–K134 and K31–K135) depends on the tertiary structure of CIFN. Based on the sequence alignments of several interferons and site-directed mutagenesis, three patches of conserved residues thought to interact with receptor(s) were found on the loop region (including the adjacent loop region) (Cys29–Phe36) of helix C and helix D (Try122–Ala139) [21–23]. It is thus reasonable to assume that the quantitative evaluation of the formation of the intra-molecular cross-links may provide very useful information on the biological activity of CIFN. It has been generally accepted that these flexible regions, such as the loop, are the key structures that generate biological functions. According to the protein structure database SCOP [24], interferon, along with interleukin 10, 19, and 20, belongs to the interferon/interleukin-10 family. When these structures were superimposed (Fig. 10), they were found to have a large amount of structural versatility at the loop regions, which connect 4 helix structures. It has been reported that the loop region connecting the helices in interferon-β is responsible for receptor binding [25]. Structural versatility in the loop region might be related to the different specificities of the protein activities. It is also important to note that recent research on the relationship between protein function and tertiary structures has indicated that there are many eukaryotic proteins that exist as disordered structures under physiological conditions, but fold into an ordered structure upon binding to their cellular targets. These proteins are called “intrinsically disordered proteins” and the phenomenon is called “coupled folding and binding” [26,27]. According to previous research results, it is reasonable that flexible and disordered structures, like a loop, are more important for activity expression than those that are rigid and ordered, such as helix, sheet, and turn structures.

In this study, the relative ratios of all cross-linked peptides were reduced to varying degrees, depending on the conditions used for denaturation, which suggests that the analytical method may be capable of quantitative evaluation. While reproducibility is an essential requirement for the quantification of cross-linked peptides, quantification using the MRM method has generally been difficult for proteins. In order to cope with this difficulty, an internal standard peptide was used for quantification in this study, and the degrees of denaturation for CIFN appeared to agree well with the results of quantification (Table 2). Interestingly, the results shown in Table 2 indicate that intra-molecular cross-linked peptides, such as K31–K134, K31–K135, and K134–135, were more sensitive to the tertiary structural changes. In contrast, formation of the inter-molecular cross-linked peptides, such as K31–K122 and K122–K134, was less sensitive to the denaturing conditions, i.e. tertiary structural changes. The relative insensitivity of the inter-molecular cross-links could be because the amino acid residues, such K121 and K122, were located on a rigid helix on the surface of the protein, making them less vulnerable to protein denaturation. Because the intra-molecular cross-links were more sensitive to the storage conditions, the analytical method promises to be all the more useful as an identification test for the tertiary protein structures that determine biological activity.

## 4. Conclusion

In this study, a relatively simple analytical method sensitive to tertiary protein structure is proposed for complementary quality testing. The analytical method is based on the use of cross-linking modifications and peptide mapping coupled with LC–MS/MS analysis. Using consensus interferon as a model protein, this study successfully identified and quantified the cross-linked peptides,

which suggests that the analytical method may be used as a tool for identifying protein pharmaceuticals. This method may be applicable to both the quantitative and qualitative evaluation of protein denaturation, and thus protein biological activity. Importantly, sensitivity to denaturing conditions during cross-linking was greater in the flexible loop region responsible for biological activity. The theoretical tertiary structure calculated using *in silico* techniques was found to be highly informative when evaluating experimental data. This method allows for easy sample preparation and requires only a small amount of test sample due to its high sensitivity, which indicates potential applicability to extra quality analysis such as process development and formulation development study.

## Acknowledgements

We are grateful to Dr. Atsuhide Mori and Dr. Gayathri Ratnaswamy for critically reading the manuscript. We also thank our Astellas colleagues, Dr. Satoshi Kitamura, Yoshimi Murata and Tatsuya Kojima, for their suggestions and support.

## References

[1] A.A. Wakankar, J. Liu, D. Vandervelde, Y.J. Wang, S.J. Shire, R.T. Borchardt, The effect of cosolutes on the isomerization of aspartic acid residues and conformational stability in a monoclonal antibody, J. Pharm. Sci. 96 (2007) 1708–1718.
[2] P.E. Wright, What can two-dimensional NMR tell us about proteins? Trends Biochem. Sci. 14 (1989) 255–260.
[3] T.L. Blundell, M. Brunori, B. Curti, B. Martino, A. Coda, M. Fumagalli, L. Ungaretti, Crystallization and preliminary X-ray diffraction studies on met-myoglobin from Aplysia limacina, J. Mol. Biol. 97 (1975) 665–666.
[4] D. Ejima, K. Tsumoto, H. Fukada, R. Yumioka, K. Nagase, T. Arakawa, J.S. Philo, Effects of acid exposure on the conformation, stability, and aggregation of monoclonal antibodies, Proteins 66 (2007) 954–962.
[5] B.T. Ruotolo, C.C. Tate, D.H. Russell, Ion mobility-mass spectrometry applied to cyclic peptide analysis: conformational preferences of gramicidin S and linear analogs in the gas phase, J. Am. Soc. Mass Spectrom. 15 (2004) 870–878.
[6] Y.F. Wang, M.Y. Ho, Y.P. Ho, Using mass spectrometry to probe the subtle differences in conformations of several cytochromes c in aqueous and methanol solutions, J. Mass Spectrom. 39 (2004) 1523–1530.
[7] Y. Li, T.D. Williams, E.M. Topp, Effects of excipients on protein conformation in lyophilized solids by hydrogen/deuterium exchange mass spectrometry, Pharm. Res. (2007) 259–267.
[8] A. Sinz, Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions, Mass Spectrom. Rev. 25 (2006) 663–682.
[9] B.X. Huang, H.-Y. Kim, C. Dass, Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry, J. Am. Soc. Mass Spectrom. 15 (2004) 1237–1247.
[10] J.W. Back, V. Notenboom, L.J. de Koning, A.O. Muijsers, T.K. Sixma, C.G. de Koster, L. de Jong, Identification of cross-linked peptides for protein interaction studies using mass spectrometry and 18O labeling, Anal. Chem. 74 (2002) 4417–4422.
[11] J.W. Back, A.F. Hartog, H.L. Dekker, A.O. Muijsers, L.J. de Koning, L. de Jong, A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes, J. Am. Soc. Mass Spectrom. 12 (2001) 222–227.
[12] M.M. Young, N. Tang, J.C. Hempel, C.M. Oshiro, E.W. Taylor, I.D. Kuntz, B.W. Gibson, G. Dollinger, High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5802–5806.
[13] G.H. Dihazi, A. Sinz, Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry, Rapid Commun. Mass Spectrom. 17 (2003) 2005–2014.
[14] B. Schilling, R.H. Row, B.W. Gibson, X. Guo, M.M. Young, MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides, J. Am. Soc. Mass Spectrom. 14 (2003) 834–850.
[15] W. Klaus, B. Gsell, A.M. Labhardt, B. Wipf, H. Senn, The three-dimensional high resolution structure of human interferon α-2a determined by heteronuclear NMR spectroscopy in solution, J. Mol. Biol. 274 (1997) 661–675.
[16] M.G. Williams, H. Shirai, J. Shi, H.G. Nagendra, J. Mueller, K. Mizuguchi, R.N. Miguel, S.C. Lovell, C.A. Innis, C.M. Deane, L. Chen, N. Campillo, D.F. Burke, T.L. Blundell, P.I. de Bakker, Sequence-structure homology recognition by iterative alignment refinement and comparative modeling, Proteins 5 (2001) 92–97.
[17] K. Mizuguchi, C.M. Deane, T.L. Blundell, M.S. Johnson, J.P. Overington, JOY: protein sequence-structure representation and analysis, Bioinformatics 14 (1998) 617–623.
[18] G.H. Kruppa, J. Schoeniger, M.M. Young, A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry, Rapid Commun. Mass Spectrom. 17 (2003) 155–162.
[19] P. Novak, M.M. Young, J.S. Schoeniger, G.H. Kruppa, A top-down approach to protein structure studies using chemical cross-linking and Fourier transform mass spectrometry, Eur. J. Mass Spectrom. 9 (2003) 623–631 (Chichester, Eng).
[20] R. Radhakrishnan, L.J. Walter, A. Hruza, P. Reichert, P.P. Trotta, T.L. Nagabhushan, M.R. Walter, Zinc mediated dimer of human interferon-alpha 2b revealed by X-ray crystallography, Structure 4 (1996) 1453–1463.
[21] E.N. Fish, Definition of receptor binding domains in interferon-α, J. Interferon Res. 12 (1992) 257–266.
[22] G.J. Waine, M.J. Tymms, E.R. Brandt, B.F. Cheetham, A.W. Linnane, Structure-function study of the region encompassing residues 26–40 of human interferon-α 4: identification of residues important for antiviral and antiproliferative activities, J. Interferon Res. 12 (1992) 43–48.
[23] Y. Mitsui, T. Senda, T. Shimazu, S. Matsuda, J. Utsumi, Structural, functional and evolutionary implications of the three-dimensional crystal structure of murine interferon-β, Pharmacol. Ther. 58 (1993) 93–132.
[24] A. Andreeva, D. Howorth, S.E. Brenner, T.J. Hubbard, C. Chothia, A.G. Murzin, SCOP database in 2004: refinements integrate structure and sequence family data, Nucleic Acids Res. 32 (2004) D226–229.
[25] T. Senda, T. Shimazu, S. Matsuda, G. Kawano, H. Shimizu, K.T. Nakamura, Y. Mitsui, Three-dimensional crystal structure of recombinant murine interferon-β, Embo. J. 11 (1992) 3193–3201.
[26] H.J. Dyson, P.E. Wright, Intrinsically unstructured proteins and their functions, Nat. Rev. Mol. Cell Biol. 6 (2005) 197–208.
[27] V.N. Uversky, Natively unfolded proteins: a point where biology waits for physics, Protein Sci. 11 (2002) 739–756.