
Probing Three-Dimensional Structure of Bovine Serum Albumin by Chemical Cross-Linking and Mass Spectrometry

Bill X. Huang and Hee-Yong Kim

Section of Mass Spectrometry, Laboratory of Membrane Biophysics and Biochemistry, NIAAA, NIH, Bethesda, Maryland, USA

Chhabil Dass

Chemistry Department, The University of Memphis, Memphis, Tennessee, USA

Serum albumin is the principal transporter of fatty acids that are otherwise insoluble in circulating plasma. While the crystal structure of human serum albumin (HSA) as well as its binding with fatty acids has been characterized, the three dimensional structure of bovine serum albumin (BSA) has not been determined although both albumins share 76% sequence homology. In this study we used mass spectrometry coupled with chemical cross-linking, to probe the tertiary structure of BSA. BSA was modified with lysine specific cross-linkers, bis(sulfosuccinimidyl) suberate (BS³), disuccinimidyl suberate (DSS) or disuccinimidyl glutarate (DSG), digested with trypsin and analyzed by tandem mass spectrometry. With O-18 labeling during the digestion, through-space cross-linked peptides were readily identified in mass spectra by a characteristic 8 Da shift. From the cross-linked peptides identified in this study, we found that 12 pairs of lysine residues were separated within 20 Å, while 5 pairs were spaced between 20 and 24 Å. The spatial distance constraints generated from five K-K pairs in BSA were consistent with the corresponding distance obtained from the crystal structure of HSA, although only six equivalent K-K pairs could be compared. According to our data, the distance between K235 of IIA and K374 of IIB domain in BSA was farther by 7–11 Å than that expected from the crystal structure of HSA, suggesting structural differences between BSA and HSA in this region. The distance constraints obtained for lysine residues using various cross-linkers should be valuable in assisting the determination of the 3-D structure of BSA. (J Am Soc Mass Spectrom 2004, 15, 1237–1247) © 2004 American Society for Mass Spectrometry

Serum albumin is the most abundant protein in the circulatory system. The most important physiological function of serum albumin is to maintain the osmotic pressure and pH of blood [1, 2], and transport a wide variety of endogenous and exogenous compounds including fatty acids, metal, amino acids, steroids and drugs [3]. Because of these extraordinary characteristics, albumins from various sources have gained extensive biomedical and industrial applications as well as research interest. The crystal structure of human serum albumin (HSA) has been determined [4, 5] and the interaction with fatty acids has been studied for HSA [3–9] and bovine serum albumin (BSA) [10–13]. Although the conformation of BSA is thought to be similar to HSA due to 76% of amino acid sequence

homology, the three-dimensional (3-D) structure of BSA has yet to be characterized.

The biological function of a protein depends on its conformation. The most informative method to study the 3-D structure of proteins is X-ray crystallography. The X-ray crystallography, which provides information regarding the atomic distances of a crystallized compound, is not always applicable because not all proteins can be readily crystallized. In addition, proteins rarely exist as crystals under physiological conditions, but are soluble in cytosol or associated with the membrane or other components of the cell. A native biochemical condition is often required to maintain an optimal biological activity of proteins. Nuclear magnetic resonance (NMR) spectroscopy can generate the tertiary structural information of proteins in solution, and therefore, complementing X-ray crystallography. However, both techniques require large amount of samples, imposing limitations in many biological settings.

Over the past years mass spectrometry, in conjunc-

Published online July 9, 2004

Address reprint requests to Dr. H.-Y. Kim, Section of Mass Spectrometry, Laboratory of Membrane Biophysics and Biochemistry, NIAAA, NIH, 12420 Parklawn Dr., Bethesda, MD 20892-8115, USA. E-mail: hykim@nih.gov

tion with chemical cross-linking, has emerged as a sensitive tool for probing the tertiary structure of proteins [14–19]. A protein is modified covalently by bi-functional cross-linking reagents, followed by proteolytic digestion. Cross-linked peptides are analyzed by mass spectrometry to assign the amino acid residues participating in the cross-linking. The maximum distance between the cross-linked residues is dependent on the cross-linker arm length. This information generates the intramolecular distance constraints, assisting computational method to determine the 3-D atomic arrangement of a protein. The approach is rapid and requires small amounts of proteins with no limitation on the size of the molecule. In addition, the method can be applied in native biochemical conditions, showing promises in studying conformational changes of a protein associated with its biological function. Recently, Back et al. has used the O-18 labeled proteolysis to unambiguously recognize the cross-linking between two peptide segments [20]. In the mass spectra of peptides obtained from the cross-linked proteins digested in O-18 water, the through-space cross-linked peptides were distinguished by a characteristic 8 Da mass shift, whereas the non-modified, internally cross-linked or end-capped peptides showed a 4 Da mass shift. The mass shifts rely on the complete incorporation of two O-18 atoms during proteolytic hydrolysis for each C-terminus ending with either lysine or arginine [21, 22].

In this study, we probe the conformation of BSA by mass spectrometry, combined with lysine-specific cross-linking and tryptic digestion. To generate information about the distance constraints for BSA, which contains 60 lysines as well as many hydrophobic pockets, we used three different lysine-specific cross-linkers of varying spacer arm lengths and hydrophobicity (Figure 1). Bis(sulfosuccinimidyl) suberate (BS^3), a water-soluble chemical cross-linker with a spacer arm length of 11.4 Å, was chosen to generate information of the lysine–lysine distance constraints for the lysines in the hydrophilic or surface regions. Disuccinimidyl suberate (DSS), a water-insoluble analog of BS^3 whose reactivity and spacer arm length are virtually identical to BS^3 [23], was used in an attempt to generate the cross-linking information for the lysines in the hydrophobic regions. In addition, we used disuccinimidyl glutarate (DSG), a hydrophobic cross-linker with a shorter spacer arm length (7.7 Å), to further refine the distance constraints between lysine residues in the BSA molecule. Each cross-linked peptide was recognized with the aid of O-18 labeling and characterized by the MS/MS analysis. The distance constraints obtained from the cross-linked lysine residues in BSA were compared with the distance between the corresponding lysine residues in HSA from the crystal structure, wherever such comparison was feasible.

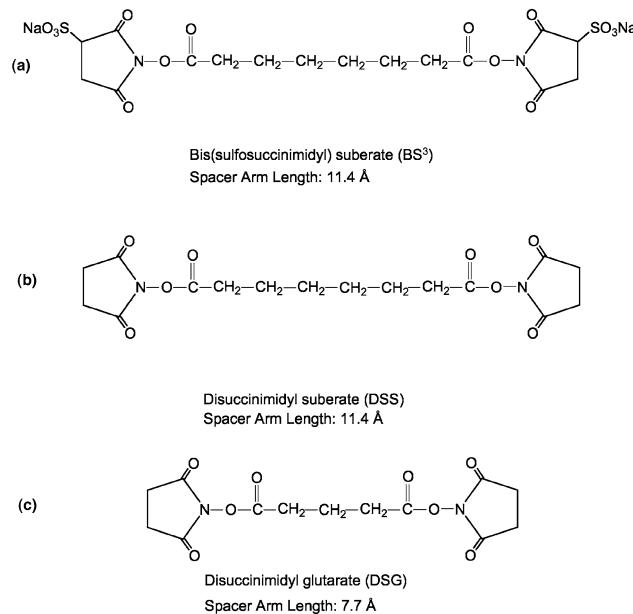


Figure 1. Molecular structures of lysine specific cross-linkers, BS^3 , DSS and DSG.

Experimental

Materials

Fatty-acid free bovine serum albumin (BSA) was purchased from Calbiochem (La Jolla, CA). Bis(sulfosuccinimidyl) suberate (BS^3), disuccinimidyl suberate (DSS), and disuccinimidyl glutarate (DSG) were purchased from Pierce (Rockford, IL). Modified trypsin was purchased from Promega (Madison, WI). O-18 enriched water (99%) was purchased from Isotec (Miamisburg, OH). Potassium phosphate was purchased from Mallinckrodt (Phillipsburg, NJ). Pure water was obtained from a Milli-Q UV plus ultra-pure water system (Bedford, MA). Other chemicals were from Sigma-Aldrich (St. Louis, MO).

Cross-Linking and Digestion

BSA at 10 μM in phosphate buffer (pH 7.8) was incubated with 500 μM cross-linkers (prepared freshly in accordance with the manufacturer's instructions) at room temperature for 30 min. This condition represents a cross-linker to protein ratio of 50:1. We have also used a cross-linker to protein ratio of 20:1, which is supposed to generate only one cross-linking per molecule [24], and found virtually identical results. The circular dichroism (CD) spectra for the BSA molecules cross-linked under our experimental condition was also identical to that of the non-cross-linked control (data not shown), indicating that no major conformational changes occurred due to cross-linking. The reaction was quenched with 1 M Tris-HCl (pH 7.4). The protein was reduced with dithiothreitol (100 mM) for 1 h at 56 °C and alkylated with iodoacetamide (100 mM) for 45 min at room temperature. The sample mixture was dialyzed

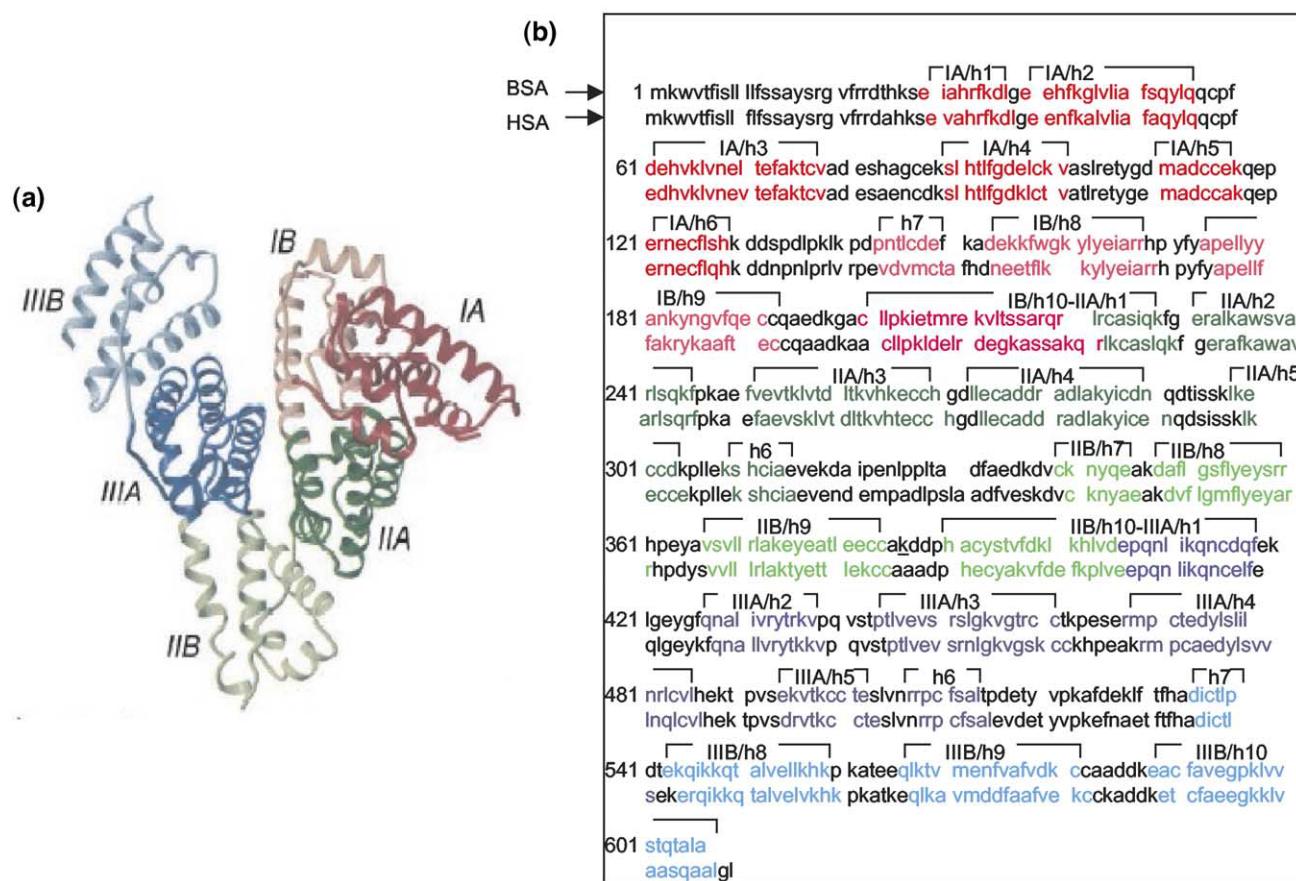


Figure 2. Domain structure and sequence homology of bovine and human serum albumin. (a) The heart-shaped serum albumin molecule consists of three homologous α -helical domains (I, II, III). Each domain contains two subdomains (A and B) that share common structural motifs [6]. (b) The homology in sequence of BSA and HSA. The homology in helical segments of each domain is shown: h1–h6 for Subdomain A and h7–h10 for Subdomain B.

against pure water using a membrane with MW cutoff of 3500 Da (Waters) three times within 24 h. The sample was divided into two equal aliquots which were dried using a vacuum centrifuge. As described by Back [20], the samples were reconstituted in either pure water or 99% O-18 enriched water containing 5% ACN. The reconstituted mixture was subjected to tryptic digestion at 37 °C for 20 h with a trypsin to protein ratio of 1:10. Alternatively, after reduction and alkylation, SDS/PAGE was carried out to isolate the intramolecular-crosslinked monomer from the intermolecular-crosslinked multimers. The monomer was subjected to in-gel tryptic digestion. Under our experimental conditions, the multimeric bands were not observed after visualization with Coomassie Brilliant Blue staining.

Off-Line Static Nano-ESI Mass Spectrometric Analysis

After desalting using a C18 ZipTip (Millipore Corp., Bedford, MA), the tryptic peptides were analyzed by a QSTAR pulsar Qq-TOF mass spectrometer (Applied

Biosystems/MDS Sciex, Toronto, CA) equipped with a nano-electrospray ionization source. The ion source voltage was set at 1100 V in the positive ion mode. A full mass spectrum was acquired over an *m/z* range of 500–2000, and the ions of interest were subjected to collision-induced dissociation using high purity argon for MS/MS analysis.

LC-Nano-ESI Mass Spectrometric Analysis

LC-nano-ESI MS analysis was performed on an Agilent 1100 series LC/MSD Trap XCT mass spectrometer (Palo Alto, CA) equipped with an Agilent 1100 nanoflow LC system. Peptides were separated using a Zorbax 300 SB C18 column (75 μ m \times 150 mm, 3.5 μ m, Agilent, Wilmington, DE) using a mobile phase that contained Solvent A (0.1% formic acid in water) and Solvent B (0.1% formic acid in ACN) at a flow rate of 300 nL/min. The mobile phase composition was held initially at 2% B for 3 min and then changed from 2–80% B over 80 min. A capillary voltage was set at 1600 V. The mass spectrometer was scanned from 300 to 2000 *m/z* range.

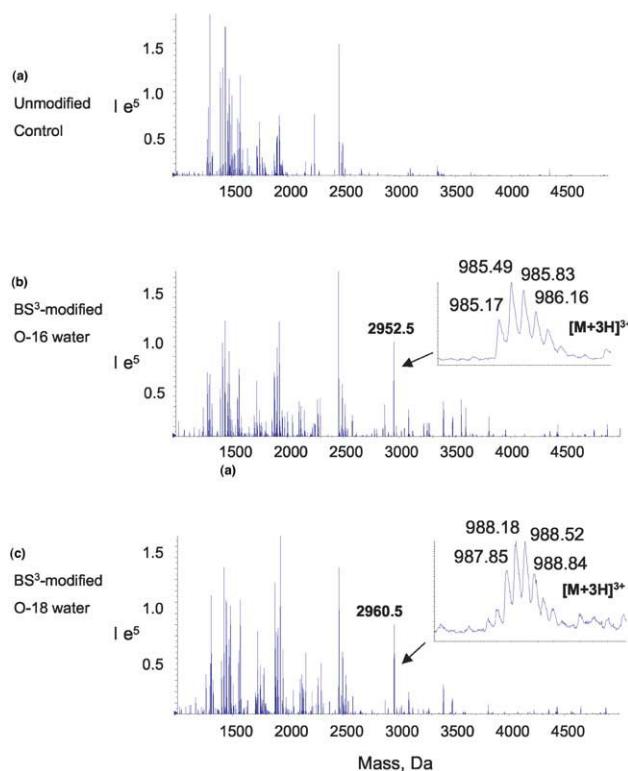


Figure 3. Nano-ESI-QqTOF reconstructed mass spectra of tryptic digests from intact (a) and BS³-modified BSA (b, c). A peak with a mass of 2952.5 Da emerged from BS³-modified BSA (b) and shifted to 2960.5 Da when digested in O-18 water (c). The shift of 8 Da indicated a through-space cross-linking between two peptide segments.

Identification of Peptides and Cross-Linked Peptides

The mass values of tryptic peptides were assigned with the help of the Protein Analysis Work Sheet (PAWS) (ProteoMetrics, NY). The cross-linked tryptic peptide peaks were distinguished by the comparison of the mass spectra obtained from the digests of the non-cross-linked control or cross-linked BSA. The through-space cross-linked peptide peaks were confirmed by the characteristic 8 Da mass shift when digested in O-18 water. The identity of the cross-linked peptides was revealed through the amino acid sequence information obtained using the MS/MS analysis.

Comparison of the 3-D Information of BSA with the Structure of HSA

The 3-D structure of BSA is believed to be very similar to HSA, because the two proteins share 76% sequence identity. Figure 2a shows the crystal structure of HSA [6] and the amino acid sequences of BSA and HSA. The heart-shaped HSA molecule is made up of three homologous α -helical domains (I, II, III) [1, 4]. Each domain contains two subdomains (A and B) that share common structure motifs. Each domain can be divided into 10 helical segments: for example, h1–h6 for the Subdomain

A and h7–h10 for the Subdomain B (Figure 2b). Human serum albumin structure is predominantly α -helical (67%) with the remaining polypeptide occurring in turns and extended or flexible regions between subdomains, and contains no β -sheets [1]. The subdomain information is indicated along with the amino acid sequences of BSA and HSA. The distance constraints between lysine residues in BSA obtained in the present study was compared with the distances of the corresponding lysine residues of HSA determined by the crystal structure, if such corresponding residues are present.

Results and Discussion

Analysis of Cross-Linked BSA Using BS³

As shown in Figure 3, the cross-linked peptides can be identified by comparing the mass spectra of tryptic digests obtained from unmodified (3a) and BS³-modified BSA (3b,c). The newly emerging peptide peaks in the spectra from BS³-modified BSA digested with trypsin in O-16 water (3b) or 99% O-18 water (3c) revealed the nature of the cross-linking. With the digestion in O-18 water, the characteristic mass shifts of 8 Da confirmed the through-space cross-linking peptides, whereas the cross-linked peptides within one segment or peptides end-capped with the cross-linker showed 4 Da shift. An example is shown in Figure 3b for a representative through-space cross-linked peptide with a mass value of 2952.5, which was reconstructed from a triply charged ion at m/z 985.2 (inset). In the spectrum of BS³-modified BSA digested in O-18 water (Figure 3c), the corresponding peak appeared at a mass of 2960.5 which was derived from the triply charged ion detected at m/z 987.9. The 8 Da difference in the mass indicated that this peak resulted from through-space cross-linking.

The identity of the cross-linked peptides detected as the triply charged ion at m/z 985.2 was determined by the amino acid sequence analyzed by MS/MS, as shown in Figure 4. Clearly, the data revealed that this peak resulted from the through-space cross-linking between the segments A[233–241]R, and L[372–386]K, with K235 from the Subdomain IIA cross-linked to K374 from the Subdomain IIB. Good coverage of y-series ions was observed for both segments. The amide bond cleavages due to the fragmentation of A[233–241]R are labeled with lower case letters, whereas those from L[372–386]K are marked with upper case letters. O-18 labeling during tryptic digestion assisted in distinguishing the C-terminal ions from the N-terminal ions. The y ions from O-18 labeled peptides showed a mass shift of 4 Da in comparison to the corresponding y ions, because these ions contained a single C-terminus. In contrast, b ions were not affected by O-18 labeling as shown for B₃ or b₃ at m/z 451 as an example. The B₃ or b₃ ion in this case included C₈H₁₂O₂ of BS³ with a net mass increase of 138 considering the loss of two hydro-

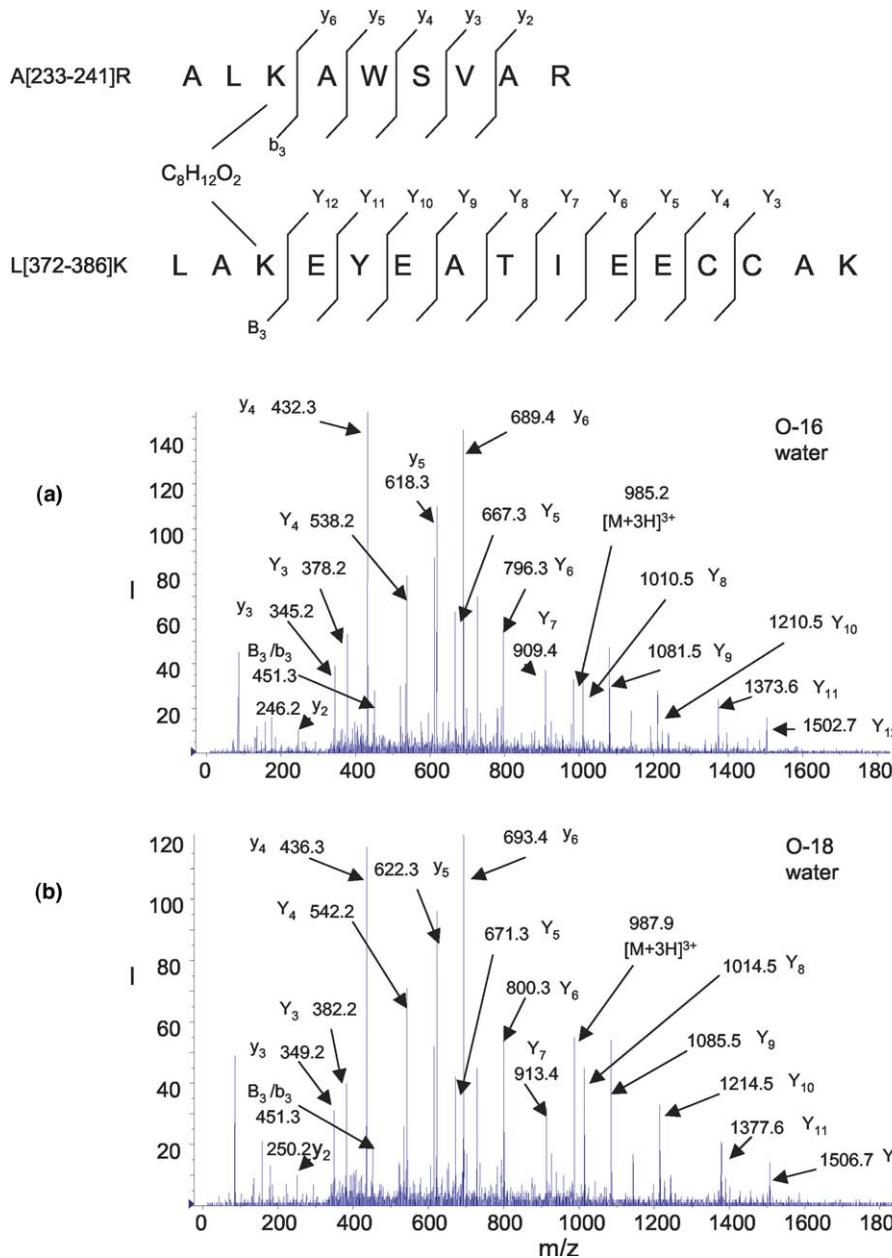


Figure 4. Nano-ESI-QqTOF-MS/MS analysis of the thru-space cross-linked peptide depicted in Figure 3. MS/MS spectra of the triply charged ion at m/z 985.2 (a) and m/z 987.9 (b) obtained from the BS³-modified BSA digested in O-16 and O-18 labeled water, respectively. The sequence of the peptide was assigned using single letter abbreviation based on the fragmentation observed in the MS/MS spectra.

gen atoms from the lysine side chains upon cross-linking.

In some cases, LC/MS/MS analyses allowed more reliable identification of cross-linked peptides as the separation of the tryptic peptide mixture precedes the mass spectrometric detection. Figure 5 depicts a typical extracted ion chromatogram (EIC) reconstructed for a cross-linked peptide with a molecular weight of 1974 Da. The peptide was detected as triply and quadruply charged ions at m/z 659.0 and 494.5, respectively, as shown in the inset of Figure 5b, and therefore, these two

ions were used to reconstruct the ion chromatogram. This chromatographic peak eluting at 27.8 min was observed in the digests of the BS³-modified BSA, and as expected, was not present in the digest from the unmodified control BSA (Figure 5a). The corresponding chromatographic peak for the O-18 labeled peptides contained the triply charged ion at m/z 661.7 and quadruply charged ion at m/z 496.5, representing a through-space cross-linked peptide with the molecular weight of 1982 (Figure 5c).

Figure 6 shows the MS/MS spectra for the triply

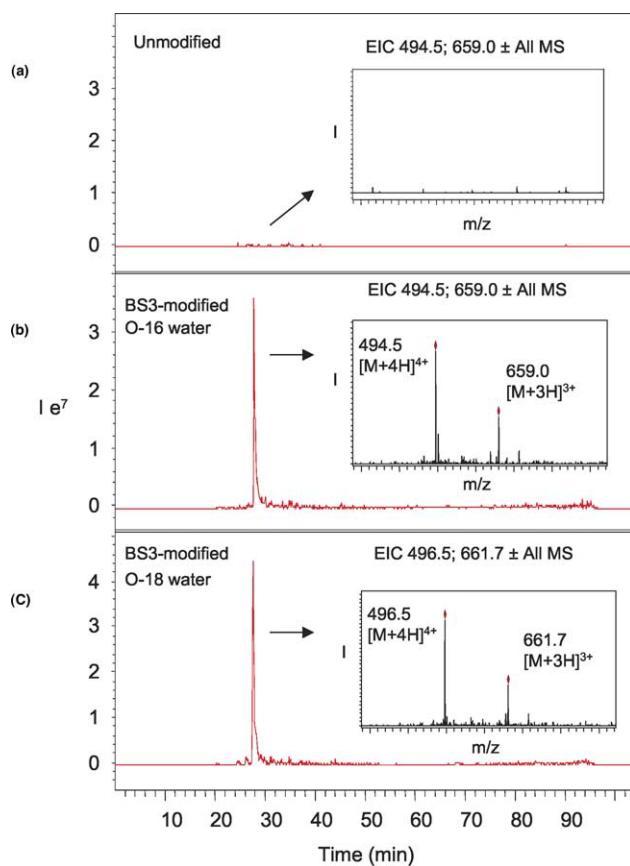


Figure 5. Extracted ion chromatograms of a cross-linked peptide obtained from unmodified (a) and BS^3 -modified BSA (b, c) analyzed by nano LC/ESI/MS/MS. A new peak with a mass of 1974, derived from the triply charged ion of m/z 659.0 and the quadruply charged ion of 494.5, eluted at 27.8 min (b). The mass shifted to 1982, derived from triply charged ion of m/z 661.7 and the quadruply charged ion of 496.5 was observed for BS^3 -modified BSA digested in O-18 water (c).

charged ions at m/z 659.0 and 661.7 derived from O-16 and O-18 labeled digests. The fragments detected in the spectra revealed that this peptide resulted from the cross-linking between the segments E[210–218]R and L[242–248]K, with K211 (located in the turn between Subdomain IB and Subdomain IIA) linked to K245 (from Subdomain IIA). The assignment of each fragment was assisted by O-18 labeling. For example, the fragment ion at m/z 1080.6 appeared in both O-16 and O-18 labeled peptide spectra, indicating that it belonged to the N-terminal ions. The singly charged ion of m/z 1225.5 increased 4 mass units in the O-18 trace, implying that this ion contained only one C-terminus. The doubly charged fragment ion of 923.0 shifted to 927.0 for the O-18 labeled peptide, reflecting a mass shift of 8 Da. Therefore, this ion was readily identified as a y-series ion containing both C-termini of the cross-linked segments. For the fragment ions containing both segments the loss of NH_3 was commonly observed as shown for y_9Y_4 , y_9Y_5 , y_9Y_6 , b_2Y_7 , and b_3Y_7 ions. The ion y_9Y_4 represented the fragment containing y_9 from the

peptide segment of E[210–218]R and Y_4 of the segment L[242–248]K plus $\text{C}_8\text{H}_{12}\text{O}_2$ of BS^3 (net mass increase of 138). According to the above information, the fragment ions detected in the MS/MS spectra of all the cross-linked peptides were unambiguously assigned. A summary of all the cross-linked and end-capped peptides detected by LC/MS and Qq-TOF is presented in Table 1. Nine through space and four internally cross-linked peptides as well as two end-capped peptides were identified and characterized. The table also lists the distances between the corresponding lysine residues from the crystal structure of HSA molecule as a comparison for the distance constraints obtained in this study.

Analysis of Cross-Linked BSA Using Hydrophobic Cross-Linkers

Figure 7 depicts representative reconstructed mass spectra of tryptic digests obtained from unmodified control (7a), DSS (7b, c) or DSG-modified BSA (7d, e). The cross-linked BSA was digested in either O-16 or O-18 water. The newly formed cross-linked peptides were identified by Qq-TOF mass spectrometric analysis of the digests from the DSS-modified or DSG-modified BSA. An example is shown for a peak with a mass of 3251.5 reconstructed from the quadruply charged ion of m/z 813.9 (inset b), which appeared in the mass spectrum of DSS-modified BSA digested with trypsin in O-16 water. Its mass shifted to 3259.6 (reconstructed from 815.9, inset c) in the corresponding O-18 spectrum, indicating a through-space cross-linking. In the mass spectrum of DSG-modified BSA digested in O-16 water, a new peak with mass of 3209.7 (reconstructed from the quadruply charged ion of m/z 803.4, inset d) was observed. The peak showed an increase of 8 Da in mass (3217.6 Da, reconstructed from 805.4, inset e) in the corresponding O-18 spectrum. The MS/MS spectrum of the peak with the mass of 3251.5 or 3259.6 (O-16 or O-18) was essentially the same as that from the peak with the mass of 3209.7 or 3217.6 (O-16 or O-18) shown in Figure 8. The MS/MS analyses revealed that the peaks originated from the cross-linking between K235 in helix 2, and K263 of helix 3 in the Subdomain IIA, via $\text{C}_8\text{H}_{12}\text{O}_2$ of DSS (net mass increase of 138) or $\text{C}_5\text{H}_6\text{O}_2$ of DSG (net mass increase of 96), respectively. The LC-MS/MS results were consistent with the data obtained from static nanospray ionization/Qq-TOF analysis. Tables 2 and 3 list all cross-linked and capped peptides that were modified by DSS and DSG, respectively. Using DSS, 12 through-space and 5 internally cross-linked peptides, and 1 end-capped cross-linked peptide were identified. With DSG, 8 through-space and 4 internally cross-linked peptides, and 1 end-capped cross-linked peptide were detected.

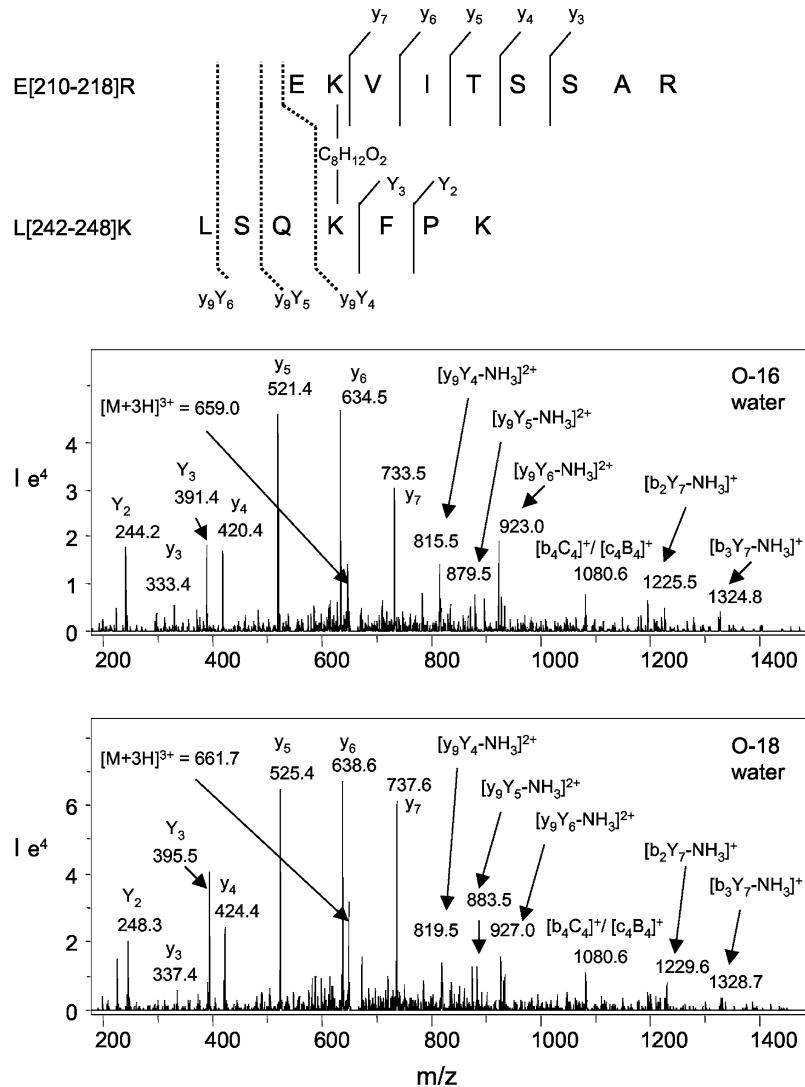


Figure 6. MS/MS spectra of the through-space cross-linked peptide depicted in Figure 5. The MS/MS spectra were obtained for the triply charged ion at m/z 659.0 (**a**) and m/z 661.7 (**b**) from the unlabeled and O-18 labeled tryptic digest, respectively. The fragments in the spectra indicated two peptide segments E[210–218]R and L[242–248]K cross-linked at K211–K245.

Distance Constraints of Lysines in BSA

The cross-linking with BS³ (Table 1) generated three short-span internal cross-linked peptides, which are within 7 amino acids in a segment (K138–K140, K156–K160, and K399–K401). These peptides were of little value in providing the 3-D spatial constraints, because BS³ coupling with two lysines could span up to 7 amino acids [14]. The remaining long-span internal cross-linking between K412 and K420 of the Subdomain IIIA, as well as the 9 through-space lysine cross-linking, should provide valuable information on the distance constraints in the 3-D structure of BSA. Six inter-domain cross-linked pairs (K140 of IA to K455 of IIIA, K228 of IIA to K489 of IIIA, K204 of IB/IIA to K495 of IIIA, K211 of IB/IIA to K463 of IIIA, K245 of IIA and K463 of IIIA, K211 of IB/IIA to K245 of IIA), 2 inter-subdomain cross-linked pairs (K437 of IIIA to K561 of

IIIB, K235 of IIA to K374 of IIB), as well as 2 intra-subdomain cross-linked peptides (K412 to K420 within IIIA, K455 to K463 within IIIA) were found.

Modification of BSA with DSS generated 17 cross-linked lysine pairs (Table 2). Of these, 11 pairs including K140–K455, K437–K561, K235–K374, K228–K489, K204–K495, K211–K463, K245–K463, K455–K463, K211–245, K412–420, and K138–K140, were seen in the BS³-modified BSA. However, three through-space cross-linked pairs including K235–K266, K235–K263, and K548–K559, along with a long-span internally cross-linked pair K88–K100, and two short-span internally cross-linked pairs, K263–K266, and K495–K498, were observed only in the DSS-modified BSA. In contrast, the internally cross-linked pairs K399–K401 and K156–K160 were observed only with the BS³-modified BSA.

When hydrophobic cross-linkers were used, 5%

Table 1. Cross-linked peptides from BS³-modified BSA

Observed mass of cross-linked peptides (O-16/O-18)	Cross-linked peptides assigned by MS/MS analysis	Cross-linked lysine residues	Type of cross-linking	Distance of corresponding lysines in HSA (Å)
3398/3406	D[131–151]K~S[452–459]R	K140(IIA)–K455(IIIA)	Thru-space	N.A. ^a
3085/3093	K[437–451]R~H[558–568]K	K437(IIIA)–K561(IIIB)	Thru-space	10.7 (K438–K562)
2953/2961	A[233–241]R~L[372–386]K	K235(IIA)–K374(IIIB)	Thru-space	13.4 (K236–K375)
2871/2879	C[223–232]R~L[483–495]K	K228(IIA)–K489(IIIA)	Thru-space	14.1 (K229–K490)
2513/2521	G[198–209]R~T[490–498]K	K204(IIA)–K495(IIIA)	Thru-space	N.A.
2293/2301	E[210–218]R~C[460–468]R	K211(IIA)–K463(IIIA)	Thru-space	N.A.
2150/2158	L[242–248]K~C[460–468]R	K245(IIA)–K463(IIIA)	Thru-space	N.A.
2120/2128	S[452–459]R~C[460–468]R	K455(IIIA)–K463(IIIA)	Thru-space	11.8 (K456–K463)
1974/1982	E[210–218]R~L[242–248]K	K211(IIA)–K245(IIA)	Thru-space	N.A.
3953/3957	H[402–433]R	K412(IIIA)–K420(IIIA)	Internal	N.A.
3220/3224	D[387–412]K	K399(IIIB)–K401(IIIB)	Internal	N.A.
2581/2585	D[131–151]K	K138(IA)–K140(IA)	Internal	N.A.
1711/1715	K[156–167]R	K156(IIB)–K160(IIB)	Internal	N.A.
3019/3023	T[76–100]K	K88(IIA)	Capping	
1702/1706	L[400–412]K	K401(IIIB)	Capping	

^aN.A.-Not Applicable

DMSO was included in the reaction mixture. However, DMSO had no influence on the integrity of the BSA conformation, as cross-linking of BSA with BS³ in the presence or absence of 5% DMSO generated virtually the same results (data not shown). One plausible explanation for the additional cross-linked pairs using DSS is that some lysine residues may not be accessible to the cross-linker BS³ due to hydrophobicity of the surrounding environments. For example, K100 is surrounded by the hydrophobic residues, including leucine, cysteine, valine, and alanine, and thus is more accessible to hydrophobic cross-linker DSS but not to BS³. This view is corroborated with the observation that the cross-linking between K88 and K100 was found only with hydrophobic cross-linker DSS and not with BS³, even though end-capping at K88 was observed with both cross-linkers. Similarly, K498 is located in the hydrophobic pockets in the IIIA domain, surrounded by cysteines. Also, K548 and K559 are located in the hydrophobic cavity of the IIIB domain; K263 and K266 are also located in the hydrophobic region in IIA [6]. These lysine residues in the hydrophobic regions may not be accessible to BS³, and therefore, were not cross-linked when a water soluble BS³ was used as a cross-linker. In contrast, K156 or K160 might be easily accessible to BS³ but not to DSS, due to the influence of the nearby hydroxyl group of Y161 [9]. As a consequence, the cross-linking of K156–K160 occurred only when BS³ was used. Likewise, K399 and K401, modified only with BS³ (Table 1), may be spatially surrounded with polar side chains, preventing the penetration of the hydrophobic cross-linkers to these lysine residues.

The lysines that are located at the interfaces between helices or turns between domains, including K138 and K140 (interface between h6 and h7 of IA); K489 (interface between h4–h5 of IIIA); K228 (right before the interface between h1 and h2 of IIA); K204 and K211 (turn between IB and IIA); K495 (near interface between

h4 and h5 of IIIA); K463 (interface between h3 and h4 of IIIA); K412 (turn between IIIB and IIIA), K420 (interface

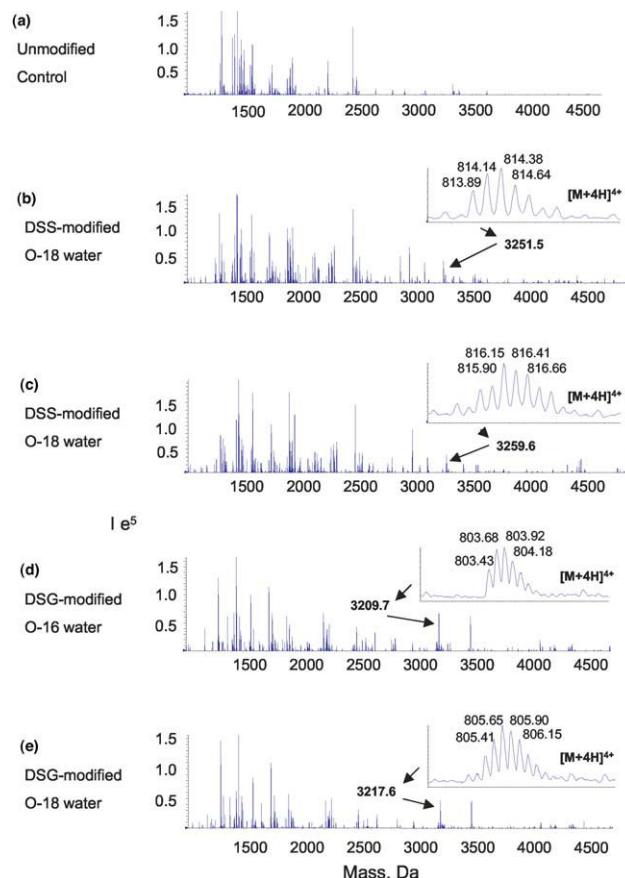


Figure 7. Nano-ESI-QqTOF reconstructed mass spectra of tryptic digests from unmodified (a) and DSS- (b, c) or DSG-modified BSA (d, e) digested in O-16 and O-18 water. The high resolution mass profile of a quadruply charged ion of a through-space cross-linked peptide observed at 3251.5 Da (DSS) or 3209.7 Da (DSG) is shown in the insets as an example.

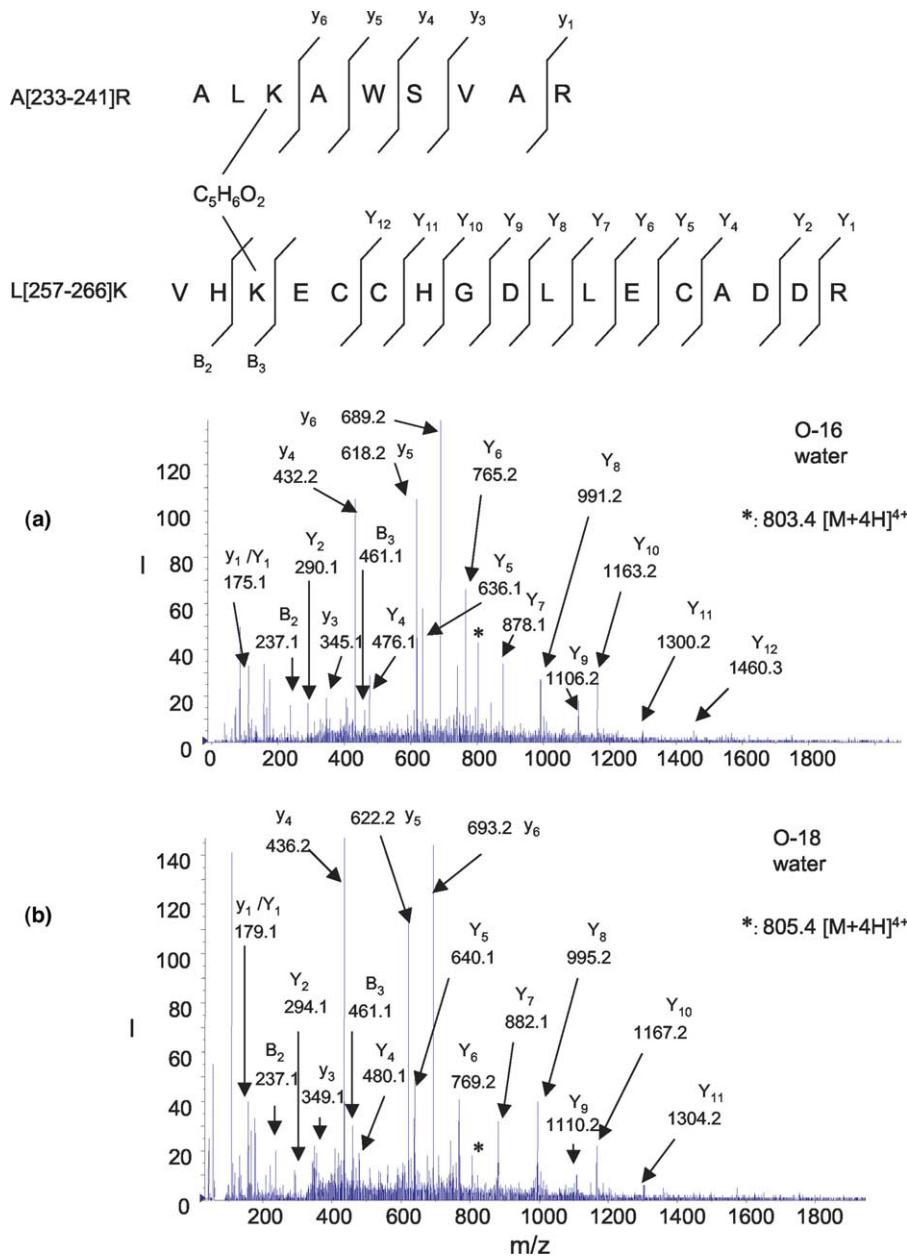


Figure 8. Nano-ESI-QqTOF-MS/MS analysis of the through-space cross-linked peptide depicted in Figure 7. The MS/MS spectra of the quadruply charged ion at m/z 803.4 (a) and m/z 805.4 (b) obtained from DSG-modified BSA digested with trypsin in unlabeled or O-18 labeled water, respectively. The structure of the peptide was assigned using single letter abbreviation, based on the fragmentation observed in the spectra. The fragments in the spectra indicated two peptide segments A[233–241]R and L[257–266]K cross-linked at K235–K263.

between h1 and h2 of IIIA); K561 (interface between h8 and h9 of IIIB); K437(near the interface between h2 and h3 of IIIA); K88 (interface between h3 and h4 of IA); K245 (near the interface between h2 and h3 of IIA), are exposed to the surface, and therefore, accessible to both water-soluble BS³ and hydrophobic DSS, for cross-linking reactions, as evidenced by most of the common cross-linked pairs. Considering the length of flexible lysine side chains, the maximum span allowed by the cross-linking with DSS or BS³ is approximately 24 Å

[14]. Therefore, the distance between the cross-linked lysine pairs should be within 24 Å.

The spacer arm length of DSG is 7.7 Å, which is shorter than that of DSS or BS³ (11.4 Å) by 3.7 Å. Accordingly, the maximum distance between lysine pairs cross-linked with DSG would be approximately 20 Å. Of the 17 cross-linked pairs found in DSS-modified BSA, only 12 were identified using DSG. This information further refined the distance constraint for these 12 lysine pairs to approximately 20 Å (Table 3). However,

Table 2. Cross-linked peptides from DSS-modified BSA

Observed mass of cross-linked peptides (O-16/O-18)	Cross-linked peptides assigned by MS/MS analysis	Cross-linked lysine residues	Type of cross-linking	Distance of corresponding lysines in HSA (Å)
3398/3406	D[131–151]K~S[452–459]R	K140(IA)–K455(IIIA)	Thru-space	N.A.*
3252/3260	A[233–241]R~V[264–280]R	K235(IIA)–K266(IIA)	Thru-space	N.A.
3085/3093	K[437–451]R~H[558–568]K	K437(IIIA)–K561(IIIB)	Thru-space	10.7 (K438–K562)
2953/2961	A[233–241]R~L[372–386]K	K235(IIA)–K374(IIIB)	Thru-space	13.4 (K236–K375)
2871/2879	C[223–232]R~L[483–495]K	K228(IIA)–K489(IIIA)	Thru-space	14.1 (K229–K490)
2513/2521	G[198–209]R~T[490–498]K	K204(IIIB)–K495(IIIA)	Thru-space	N.A.
2293/2301	E[210–218]R~C[460–468]R	K211(IIIB)–K463(IIIA)	Thru-space	N.A.
2291/2299	A[233–241]R~L[257–266]K	K235(IIA)–K263(IIA)	Thru-space	10.2 (K236–K264)
2150/2158	L[242–248]K~C[460–468]R	K245(IIA)–K463(IIIA)	Thru-space	N.A.
2120/2128	S[452–459]R~C[460–468]R	K455(IIIA)–K463(IIIA)	Thru-space	11.8 (K456–K463)
1974/1982	E[210–218]R~L[242–248]K	K211(IIIB)–K245(IIA)	Thru-space	N.A.
1788/1792	K[548–557]K~H[558–561]K	K548(IIIB)–K559(IIIB)	Thru-space	16.6 (K549–K560)
3953/3957	H[402–433]R	K412(IIIA)–K420(IIIA)	Internal	N.A.
3528/3532	T[76–105]R	K88(IA)–K100(IA)	Internal	N.A.
3021/3025	L[257–280]R	K263(IIA)–K266(IIA)	Internal	N.A.
2581/2585	D[131–151]K	K138(IA)–K140(IA)	Internal	N.A.
2245/2249	T[490–507]R	K495(IIIA)–K498(IIIA)	Internal	N.A.
3019/3023	T[76–100]K	K88(IA)	Capping	

*N.A.-Not Applicable

the remaining five cross-linked pairs, including K140–K455, K235–K374, K245–K463, K211–K245, and K412–420, were not found, implying that the K–K distances for these lysine pairs are greater than 20 Å but less than 24 Å.

Of the total 19 cross-linked lysine pairs observed with BSA, only 6 homologous pairs can be found in the HSA crystal structure (PDB entry 1ao6). The K–K distances of these pairs were obtained using Protein Explorer (version 1.982) (Tables 1–3). The distance constraints obtained in this study are consistent with the HSA crystal structure with the exception of K235–K374. The distance between K235 and K374 in BSA is expected to be between 20 and 24 Å, based on the fact that they were cross-linked with DSS but not with DSG. However, the distance between the corresponding lysine

residues K236 and K375 in HSA is 13.4 Å, suggesting that the 3-D structure of BSA may deviate from that of HSA in this region. The information on the distance constraints obtained from the cross-linked lysine residues, particularly for which homologous lysines were not found in HSA, should be valuable in assisting the characterization of BSA conformation.

Conclusions

We have used mass spectrometric analysis coupled with chemical cross-linking to probe the 3-D structure of BSA. Because BSA contains 60 lysines, lysine specific cross-linkers were employed. Considering that many lysine residues exist in the hydrophobic regions, the

Table 3. Cross-linked peptides from DSG-modified BSA

Observed mass of cross-linked peptides (O-16/O-18)	Cross-linked peptides assigned by MS/MS analysis	Cross-linked lysine residues	Type of cross-linking	Distance of corresponding lysines in HSA (Å)
3210/3218	A[233–241]R~V[264–280]R	K235(IIA)–K266(IIA)	Thru-space	N.A.*
3043/3051	K[437–451]R~H[558–568]K	K437(IIIA)–K561(IIIB)	Thru-space	10.7 (K438–K562)
2829/2837	C[223–232]R~L[483–495]K	K228(IIA)–K489(IIIA)	Thru-space	14.1 (K229–K490)
2471/2479	G[198–209]R~T[490–498]K	K204(IIIB)–K495(IIIA)	Thru-space	N.A.
2251/2259	E[210–218]R~C[460–468]R	K211(IIIB)–K463(IIIA)	Thru-space	N.A.
2249/2257	A[233–241]R~L[257–266]K	K235(IIA)–K263(IIA)	Thru-space	10.2 (K236–K264)
2078/2086	S[452–459]R~C[460–468]R	K455(IIIA)–K463(IIIA)	Thru-space	11.8 (K456–K463)
1746/1754	K[548–557]K~H[558–561]K	K548(IIIB)–K559(IIIB)	Thru-space	16.6 (K549–K560)
3486/3490	T[76–105]R	K88(IA)–K100(IA)	Internal	N.A.
2979/2983	L[257–280]R	K263(IIA)–K266(IIA)	Internal	N.A.
2539/2543	D[131–151]K	K138(IA)–K140(IA)	Internal	N.A.
2203/2207	T[490–507]R	K495(IIIA)–K498(IIIA)	Internal	N.A.
2977/2981	T[76–100]K	K88(IA)	Capping	

*N.A.-Not Applicable

hydrophobic cross-linkers DSS and DSG, as well as the water-soluble cross-linker, BS³, were used to obtain complementary information. The through-space cross-linked peptides were identified rapidly with the assistance of O-18 labeling during digestion with trypsin. The O-18 labeling also assisted in the interpretation of MS/MS spectra of the cross-linked peptides by distinguishing the C-terminal ions from N-terminal ions. Each cross-linked peptide was characterized by MS/MS for unambiguous assignment. The distance constraints were obtained from 12 through-space and 7 internally cross-linked lysine pairs identified in this study. Our data indicated that the distance between K140 of IA and K455 of IIIA, K235 of IIA and K374 of IIB, K245 of IIA and K463 of IIIA, K211 of IB/IIA and K245 of IIA, and K412 and 420 of IIIA, is between 20 and 24 Å. The distance constraints for 12 lysine pairs including K235–K266, K437–K561, K228–K489, K204–K495, K211–K463, K235–K263, K455–K463, K548–K559, K88–K100, K263–K266, K138–K140, and K495–K498, were found to be approximately 20 Å. The distance constraints obtained in this study also suggested that K235 and K374 are farther away from each other in BSA molecule in comparison to the corresponding lysine residues in HSA. We have demonstrated in this study that chemical cross-linking with mass spectrometric detection can be employed to readily generate the 3-D information of BSA. With proper choice of cross-linking criteria, for example, using multiple cross-linkers with different span lengths and chemistry, comprehensive information about the tertiary structure can be obtained for proteins of interest.

References

- Carter, D. C.; Ho, J. X. Structure of Serum Albumin. *Adv. Protein Chem.* **1994**, *45*, 153–203.
- Figge, J.; Rossing, T. H.; Fencl, V. The Role of Serum-Proteins in Acid-Base Equilibria. *J. Lab. Clin. Med.* **1991**, *117*, 453–467.
- Carter, D. C.; He, X.; Munson, S. H.; Twigg, P. D.; Gernert, K. M.; Broom, M. B.; Miller, T. Y. Three-Dimensional Structure of Human Serum Albumin. *Science* **1989**, *244*, 1195–1198.
- He, X. M.; Carter, D. C. Atomic Structure and Chemistry of Human Serum Albumin. *Nature* **1992**, *358*, 209–215.
- Sugio, S.; Kashima, A.; Mochizuki, S.; Noda, M.; Kobayashi, K. Crystal Structure of Human Serum Albumin at 2.5 Å Resolution. *Protein Eng.* **1999**, *12*, 439–446.
- Curry, S.; Mandelkow, H.; Brick, P.; Franks, N. Crystal Structure of Human Serum Albumin Complexed with Fatty Acid Reveals an Asymmetric Distribution of Binding Sites. *Nature Struct. Biol.* **1998**, *5*, 827–835.
- Bhattacharya, A. A.; Grune, T.; Curry, S. Crystallographic Analysis Reveals Modes of Binding of Medium and Long Chain Fatty Acids to Human Serum Albumin. *J. Mol. Biol.* **2000**, *303*, 721–732.
- Peptipsa, I.; Grune, T.; Bhattacharya, A. A.; Curry, S. Crystal Structure of Human Serum Albumin Complexed with Mono-unsaturated and Polyunsaturated Fatty Acids. *J. Mol. Biol.* **2001**, *314*, 955–960.
- Zunszain, P.; Ghuman, J.; Komatsu, T.; Tsuchida, E.; Curry, S. Crystal Structure Analysis of Human Serum Albumin Complexed with Hemin and Fatty Acid BMC. *Struct. Biol.* **2003**, *3*, 1–9.
- Cistola, D. P.; Small, D. M.; Hailton, J. A. Carbon 13 NMR Studies of Saturated Fatty Acids Bound to Bobine Serum Albumin, I. The Filling of Individual Fatty Acid Binding Sites. *J. Biol. Chem.* **1987**, *262*, 10971–10979.
- Cistola, D. P.; Small, D. M.; Hailton, J. A. Carbon 13 NMR Studies of Saturated fatty Acids Bound to Bobine Serum Albumin, II. Electrostatic Interactions in individual Fatty Acid Binding Sites. *J. Biol. Chem.* **1987**, *262*, 10980–10985.
- Choi, J. K.; Curry, S.; Qin, D.; Bittman, R.; Hamilton, J. A. Interactions of Very Long-chain Saturated Fatty Acids with Serum Albumin. *J. Lipid Res.* **2002**, *43*, 1000–1010.
- Hamilton, J. A.; Era, S.; Bhamidipati, S. P.; Reed, R. G. Location of the Three Primary Binding Sites for Long-chain Fatty Acids on Bovine Serum Albumin. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2051–2054.
- Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gilson, B. W.; Dollinger, G. High Throughout Protein Fold Identification by Using Experimental Constraints Derived from Intramolecular Cross-links and Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802–5806.
- Rappsiber, J.; Siniossoglou, S.; Hurt, E. C.; Mann, M. A Generic Strategy to Analyze the Spatial Organization of Multi-protein Complexes by Cross-Linking and Mass spectrometry. *Anal. Chem.* **2000**, *72*, 267–275.
- Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O. Isotope-Tagged Cross-linking Reagents. A New Tool in Mass Spectrometric Protein Interaction Analysis. *Anal. Chem.* **2001**, *73*, 1927–1934.
- Bennett, K. L.; Kussmann, M.; Bjork, P.; Godzow, M.; Mikkelsen, M.; Sorensen, P.; Roepstorff, P. Chemical Cross-linking with Thiol-Cleavable Reagents Combined with Differential Mass spectrometric Peptide Mapping—A Novel Approach to Assess Intermolecular Protein Contacts. *Protein Sci.* **2000**, *9*, 1503–1518.
- Taverner, T.; Hall, N. E.; Ohair, R. A. J.; Simpson, R. J. Characterization of an Antagonist Interleukin-6 Dimer by Stable Isotope Labeling, Cross-linking and Mass Spectrometry. *J. Biol. Chem.* **2002**, *277*, 46487–46492.
- Pearson, K. M.; Pannell, L. K.; Fales, H. M. Intramolecular Cross-linking Experiments on Cytochrome C and Ribonuclease A Using an Isotope Multiplet Method. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149–159.
- Back, J. W.; Notenboom, V.; Koning, L. J.; Muijsers, A. O.; Sixma, T. K.; Koster, C. G.; Jong, L. Identification of Cross-Linked Peptides for Protein Interaction Studies Using Mass Spectrometry and ¹⁸O Labeling. *Anal. Chem.* **2002**, *74*, 4417–4422.
- Schnolzer, M.; Jedrzejewski, P.; Lehmann, W. D. Protease-Catalyzed Incorporation of ¹⁸O Into Peptide Fragments and Its Application for Protein Sequencing by Electrospray and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Electrophoresis* **1996**, *17*, 945–953.
- Reynolds, K. J.; Yao, X.; Fenselau, C. J. Proteolytic ¹⁸O Labeling for Comparative Proteomics: Evaluation of Endoprotease Glu-C as the Catalytic Agent. *J. Proteome Res.* **2002**, *1*, 27–33.
- Haniu, M.; Narhi, L. O.; Arakawa, T.; Elliott, S.; Rohde, M. F. Recombinant human erythropoietin (rHuEPO): Cross-Linking with Disuccinimidyl Esters and Identification of the Interfacing Domains in EPO. *Protein Sci.* **1993**, *2*, 1441–1451.