Strategy for Selective Chemical Cross-Linking of Tyrosine and Lysine Residues

Michael D. Leavell, Petr Novak, Christopher R. Behrens,
Joseph S. Schoeniger, and Gary H. Kruppa
Sandia National Laboratories, Livermore, California, USA

Chemical cross-linking of proteins combined with mass spectral analysis is a powerful technique that can be utilized to yield protein structural information, such as the spatial arrangement of multi-protein complexes or the folding of monomeric proteins. The succinimidyl ester cross-linking reagents are commonly used to cross-link primary amine-containing amino acids (N-terminus and lysine). However, in this study they were used to react with tyrosines as well, which allowed for the formation of cross-links between two primary amines, one primary amine and one tyrosine, or two tyrosines. This result is extremely important to the chemical cross-linking community for two reasons: (1) all possible cross-linked residues must be considered when analyzing data from these experiments to generate correct distance constraints and structural information, and (2) utilizing the versatility of these cross-linking reagents allows more information content to be generated from a single cross-linking reagent, which may increase the number of cross-links obtained in the experiment. Herein, we study the reactivity of the succinimidyl ester labeling and cross-linking reagents with angiotensin I and oxidized insulin β-chain. Using the succinimidyl acetate labeling reagent, the reactivity of the N-terminus was found to be greater than either lysine or tyrosine. However, a selectivity of the cross-linking reagent was observed for either tyrosine or lysine depending on the pH of the reaction solution. In acidic pH, it was observed that tyrosine was more reactive, while in alkaline pH lysine was more reactive. Exploiting this selectivity predominantly N-terminus-tyrosine or tyrosine-tyrosine cross-links were favored at acidic pH, while N-terminus-tyrosine or tyrosine-lysine cross-links were favored at alkaline pH.   (J Am Soc Mass Spectrom 2004, 15, 1604–1611) © 2004 American Society for Mass Spectrometry

Inter-molecular chemical cross-linking has a long history as a tool for the study of the quaternary structure of protein complexes [1, 2], and many years ago it was suggested that intra-molecular cross-linking could be used as a method of obtaining distance constraints that would be useful in developing structural models of proteins [3]. However, the promise of intra-molecular cross-linking for structural modeling has only recently been enabled by state of the art mass spectrometric methods that make determination of the cross-linked residues in a protein practical on a reasonable time scale and with small quantities of protein. The first study to derive a sufficient number of intra-molecular distance constraints to develop a structural model for a protein used fibroblast growth factor 2 (FGF-2) as a test case [4]. FGF-2 is a 17 kDa protein, which has 14 lysines evenly dispersed in its 155 residue sequence. Using only the commercially available primary amine reactive cross-linker bis(sulfosuccinimidyl) suberate, 18 cross-links were found in FGF-2 in this study, of which 15 provided useful distance constraints for deter-mining the fold family of the protein. Using these 15 constraints it was possible to assign the FGF2 to the correct β–trefoil fold family, from among a large set of structures predicted using threading techniques. The name MS3D has been proposed for the study of protein 3-dimensional structure using chemical cross-linking and mass spectrometry [5].

All of the subsequent recent work published on intra-molecular cross-linking for MS3D studies has employed similar, commercially available primary amine selective reagents that can cross-link lysine residues to other lysine residues or the amino terminus [6–10]. In our own recent work, we showed that a top down approach to the localization of cross-links using FTMS could be applied to localize three cross-links in ubiquitin [7], and in a more recent publication we showed that the distance constraints can be improved by using a series of cross-linkers of different lengths [6]. However, ubiquitin has eight primary amino groups (seven on lysines and one at the amino terminus) that should be reactive with dissuccinimidyl esters, the reagents that were used in the preliminary studies described above, yet only three cross-links were observed. Examination of the known structure of ubiquitin showed that there are many geometrically allowed cross-links be-

tween primary amino groups that were not observed experimentally. To determine whether our method could not localize these cross-links or if the cross-links were not formed, we probed the reactivity of the primary amino groups of ubiquitin by reaction with N-hydroxysuccinimidyl acetate (NHSAc), which specifically acetylates primary amines by the elimination of N-hydroxy succinimide; this is the same elimination reaction that occurs twice in the cross-linking reaction with the disuccinimidyl esters [11]. As the stoichiometric ratio of NHSAc:protein increased, identification of the fragments from native protein and protein with successively increasing modification allowed the assignment of the complete order of reactivity of the primary amino groups in ubiquitin (Met1 ≈ Lys 6 ≈ Lys 48 ≈ Lys 63 > Lys 33 > Lys11 > Lys 27, Lys 29). These results are in excellent agreement with the reactivity expected from other studies and predicted from the known crystal structure of ubiquitin, which shows that the least reactive lysines are involved in strong hydrogen bonding interactions. The reactivity order observed also correlates well with the observed cross-links Lys 6-Lys 11, Met 1-Lys 6, and Lys 48-Lys 63, which all involve one of the most reactive primary amines with another nearby, reactive primary amine. Clearly the number of cross-links observed when starting from the singly cross-linked parent ion is limited by the lysine reactivity.

One could hope to overcome this limitation by starting with parent ions containing multiple cross-links, but the abundance of multiply cross-linked species is often quite low, and heavily modifying the protein may perturb its native state structure. In a recent study by another group using proteolytic digests to localize the cross-links in cytochrome-*c*, only one cross-link was observed with lysine reactive cross-linkers when the reaction was optimized to yield a single cross-link per protein molecule [9]. At higher cross-linker concentration optimized to yield two cross-links per protein molecule, the digest yielded only four additional cross-links. Only two of these five total cross-links in cytochrome-*c* were structurally significant, the others were between lysines that were very close together in the sequence and would have been expected to cross-link. This study combined with our own work, and discussions with others working in the area, provides support for our hypothesis that in order to obtain a large number (more than one cross-link/ kDa) of structurally significant cross-links it is necessary to take advantage of reactive side-chains other than lysine. There are several other advantages to examining other reactive sidechains, one is that these may lead to better Cα-Cα distance constraints: the lysine side chain is quite long and flexible, about 6 Å from Cα to Nε, so a cross-link between two lysines gives a minimum 16 Å Cα-Cα distance constraint even with a short cross-linker arm containing two methylene groups. The distribution of lysines in a protein can also be problematic, e.g., the cytochrome-*c* sequence contains one set of three

adjacent lysines, and three pairs of adjacent lysines. Cross-links at these positions would be hard to resolve by proteolytic digestion and/or MS/MS, and cross-links between lysines that are adjacent in the sequence provide no new structural information. Having a variety of cross-linkers with different sidechain specificities to choose from allows a cross-linking regimen to be chosen for a given protein based on the abundance and sequence distribution of its reactive amino acid sidechains, improving the information yield and making resolution of the cross-links via a top down MS/MS approach and/or proteolytic digestion much easier. We present here a method for cross-linking tyrosine residues through the use of the succinimidyl ester cross-linkers.

The reaction of the succinimidyl ester cross-linkers was studied with angiotensin I and oxidized insulin-β-chain. Normally, ubiquitin is our protein of choice for cross-linking methods development. However, ubiquitin contains only one tyrosine that is involved in a hydrogen bond to the protein backbone, and is therefore unreactive. Using angiotensin I and oxidized insulin-β-chain we show that the succinimidyl ester cross-linkers can be used to form cross-links not only between lysines, but also between lysine and tyrosine and even between tyrosines. In the early development and characterization of these cross-linkers [12–14], the reactivity towards tyrosine and other amino acids (e.g., histidine, cystine, and serine) was noted, but has largely been overlooked in recent studies. The reactivity of histidine and serine can generally be ignored owing to the instability of the N-acylimidazole product (histidine) and the slow reactivity of serine. While cystine was shown to be reactive, it is generally involved in disulfide bonds and therefore does not react with the succinimidyl esters. Consideration of the dual reactivity with tyrosine is paramount to the correct interpretation of cross-linking data. Exploiting the reactivity of these cross-linkers with both tyrosine and lysine should increase the number of cross-links that can be formed in a given protein and fits into our strategy of developing a diverse set of cross-linker chemistries to cross-link all of the reactive amino acid residues.

Experimental

Sample Preparation

Labeling and cross-linking reactions were carried out by first dissolving the peptide (insulin β-chain oxidized or angiotensin I) in the appropriate buffer at a concentration of 0.5 mg/mL. For experiments performed at pH = 6.0, peptides were dissolved in a pyridine/HCl buffer (50 mM), while at pH = 8.4 peptides were dissolved in a triethylamine/carbonate buffer (50 mM). The labeling reagents were dissolved in dry DMSO before reaction at a concentration of 10 mg/mL (the high concentration was chosen to minimize the concentration of DMSO in the final reaction mixture to avoid

denaturing the protein, only a few microliters of the cross-linker solution in DMSO were added to ca. 100 μl of the protein in buffer). The peptides and labeling or cross-linking reagents were combined to yield a 1:3 or 1:9 molar ratio of peptide to reagent keeping the peptide concentration at 1 mM (typical reaction volumes were ca. 100 μL). The samples were reacted for 2 h before being purified on a macro-trap (Macro-traps, Michrom BioResources, Auburn, CA) to remove salts, buffers and hydrolyzed cross-linker.

Sample Analysis

Samples were analyzed on a Bruker Apex II Q-FTMS (Bruker Daltonics, Billerica, MA) equipped with a 7.0 T superconducting magnet. This instrument was described in detail in a previous publication [11]. After labeling or cross-linking and purification, protein samples were diluted to 10 μM in 50:50 MeOH:H_2O with 6% acetic acid and were introduced via direct infusion to an Apollo ESI ion source (Bruker Daltonics, Billerica, MA) at a flow rate of 1.5 μL/min. Ions were generated in the positive ion mode. Mass spectra were obtained by accumulating the ions in the ESI source hexapole before being transferred through the Q-front end to the ICR cell for analysis. For tandem mass spectrometric analysis ions were passed through the ESI source hexapole, the ion of interest was isolated (5 Da isolation window) and fragmented in the collision cell, and the fragments were accumulated for 5 s before being transferred to the ICR cell for analysis. Tandem mass spectra were acquired over the range from 200–2500 Da. Collision energies in the Q-front end were increased until extensive fragmentation was observed. Spectra are composed of either 256 or 512 k data points, and are an average of 64 or 128 scans. Typically each MS/MS spectrum took ~5 min to acquire.

Chemicals and Materials

Angiotensin I, oxidized Insulin β-chain, pyridine, and triethylamine were obtained from Sigma-Aldrich Chemical Company (St. Louis, MO). Succinimidyl acetate (NHSAc) and ethylene glycol *bis*[succinimidylsuccinate] (EGS) were obtained from Pierce (Rockford, IL). All solvents were of HR-GC grade and were supplied by EM Science (Gibbstown, NJ).

Results and Discussion

Reactivity of N-hydroxysuccinimidyl Acetate

Succinimidyl esters are commonly utilized to cross-link primary amines (ε-amine on lysine and the α-amine on the *N*-terminus). However, in the early bio-chemical characterization of these active esters as cross-linking reagents, it was found that they showed varying reactivity with the following residues: histidine imidazole > primary amino groups $\gg$ cysteine thiolate $\approx$ tyrosine

phenolate > serine [12–14]. Therefore, it is possible that these reagents can be used to cross-link amino acids other than the N-terminus and lysine. In using these cross-linkers to obtain structural information, consideration of all possible reactive sites is paramount to derive the correct distance constraints from the mass spectral data on cross-linked proteins. Furthermore, utilizing the versatility of these reagents would allow more structural information to be gained. This is especially true if the specificity of the reagent can be adjusted based on the reaction conditions. In considering the possible cross-links formed by the succinimidyl esters, the following three points must be considered: it was found previously that the *N*-acylimidazole product (formed when His is modified) is transient and hydrolyses quickly, cystine is commonly unreactive as it is often involved in disulfide bonds, and the kinetics of serine reactivity are quite slow and cannot compete with the other fast reactions (e.g., with N-terminus, lysine, and tyrosine) [12–15]. Therefore, the primary amino group-containing amino acids and tyrosine are the most likely amino acids to form stable cross-links with these esters in soluble proteins. However, if it is known that free cysteins are contained within the protein, as is commonly observed in membrane proteins, these must be considered as well.

To investigate the reactivity of these reagents with tyrosine a simple succinimidyl ester, N-hydroxysuccinimidyl acetate (NHSAc), was mixed with angiotensin I. The reaction was run at both pH = 6.0 and pH = 8.4; identical results were obtained for both pHs. Angiotensin I is a decapeptide containing several reactive residues, including one tyrosine, an α-amine (at the N-terminus) and two histidines. The mass spectrum of acetylated angiotensin I is shown in Figure 1a; the reaction scheme is shown in Scheme **1**. From the mass spectrum, it is clear that up to two sites are modified on this peptide (Figure 1a), this was observed for both the 1:3 and 1:9 ratios of angiotensin I:NHSAc studied. MS/MS was performed to localize the site(s) of acetylation. Figure 1b shows the region around the b_3 ion when MS/MS is performed on the *singly-acetylated* precursor (*m/z* 669.86) at pH = 6.0. Since only the acetylated b_3 ion is observed, and not its unmodified counterpart, it is evident that the Asp 1 (the N-terminus) reacts completely. As shown in Figure 1c the (2-5) internal ion is observed only without modification in the MS/MS spectrum of the singly-acetylated parent ion, showing that Tyr 4 is not modified. However, in the MS/MS spectrum of the *doubly-modified* angiotensin I (*m/z* 690.86), it is clear that both Tyr 4 and the Asp 1 are modified (Figure 1d and e). The modification of Tyr 4 is shown by the absence of the unmodified (2-5) internal ion, and the presence of the (2-5)+Ac internal ion (Figure 1e). Thus, it is apparent that the succinimidyl esters are selective for the *N*-terminus over tyrosine, but do react with tyrosine, as was observed in a previous study [12–14]. In these same studies, it was also shown that histidine could be transiently modified by the
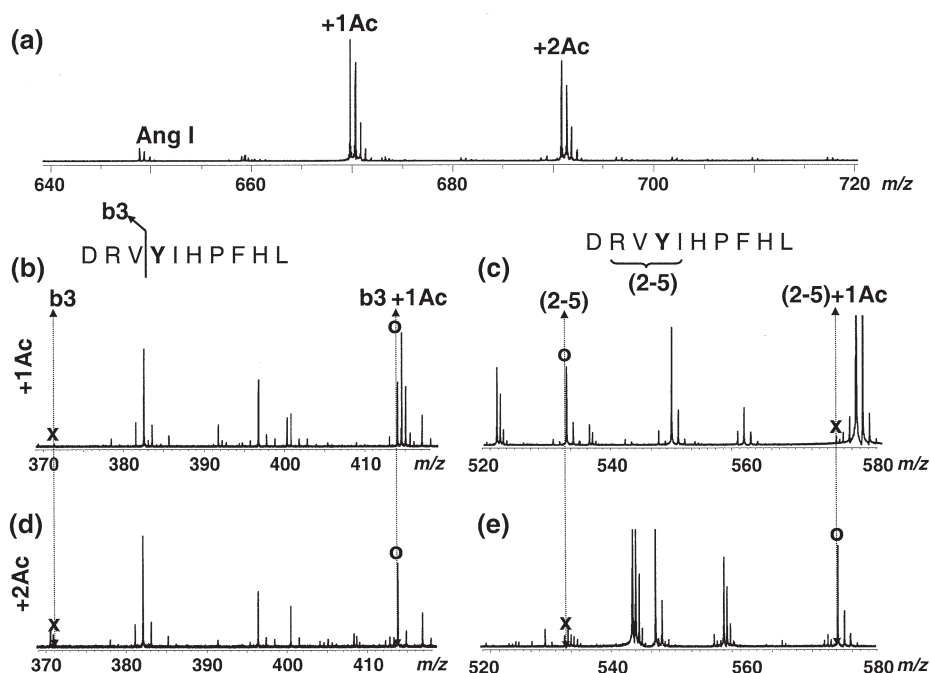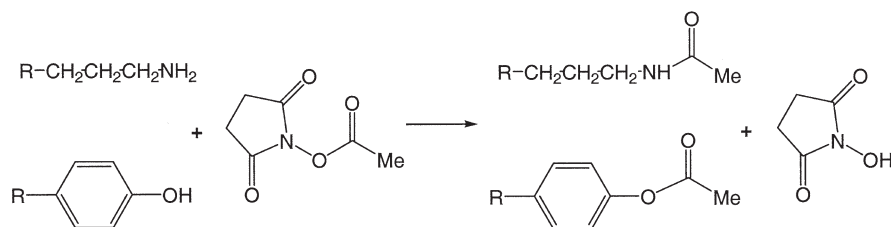
Figure 1. (**a**) Mass spectrum of the reaction of N-hydroxysuccinimidyl acetate (NHSAc) with oxidized Angiotensin I (1:3 ratio), reacted at pH = 6.0. (**b**, **c**) Product ion spectrum of singly-modified Angiotensin I expanded around the b3 (**b**) and the internal ions (**c**). (**d**, **e**) MS/MS of doubly-modified Angiotensin I expanded around the b3 ion (**d**) and the (2–5) internal ion (**e**). Note, all spectra are multiply charged, an "**x**" indicates the absence of an ion, while an "**o**" indicates the presence of an ion, and in panel (**b**), the doubly-charged b6 + Ac ion is observed, along with singly-charged b3 + Ac. In the doubly-modified spectra (**d**) this overlap does not occur as the b6 ion is doubly-modified and the b3 ion is singly-modified.

succinimidyl esters [12, 13]. Our results show that the histidine imidazole is not stable on the timescale of our experimental protocol (~2 h) and instead is quickly hydrolyzed, owing to the rapid hydrolysis of the N-acylimidazole adduct. This was indicated from the analysis of ions such as y3 or y5, neither of which was observed as acetylated species (data not shown). Therefore, adducts with histidine do not need to be considered when cross-linking with the succinimidyl ester reagents.

With the information in hand that the succinimidyl esters form stable products with tyrosines, the next step was to investigate whether these reagents can form cross-links between primary amines and tyrosine, and between two tyrosines. To probe this possibility, oxidized insulin β-chain was reacted with NHSAc, yielding the mass spectrum shown in Figure 2. In oxidized

insulin-β-chain cysteine is oxidized to a sulfonate and is therefore unreactive with the succinimidyl ester. Thus, the only amino acids available that are reactive on oxidized insulin β-chain are the primary amine-containing amino acids and tyrosine. From this spectrum it is apparent that NHSAc reacts up to 4 times with oxidized insulin-β-chain, which is consistent with the total number of primary amines (2) and tyrosines (2). The number of modifications observed was the same for both pHs studied as well as the different ratios of insulin-β-chain: NHSAc studied.

To determine the optimum reaction conditions for tyrosine-tyrosine cross-linking with succinimidyl esters, further investigation of the reaction of oxidized insulin β-chain with NHSAc was undertaken at pH 6.0 and 8.4 to determine the reactive residues at each pH and the order of reactivity of each residue. In order to

Scheme 1.

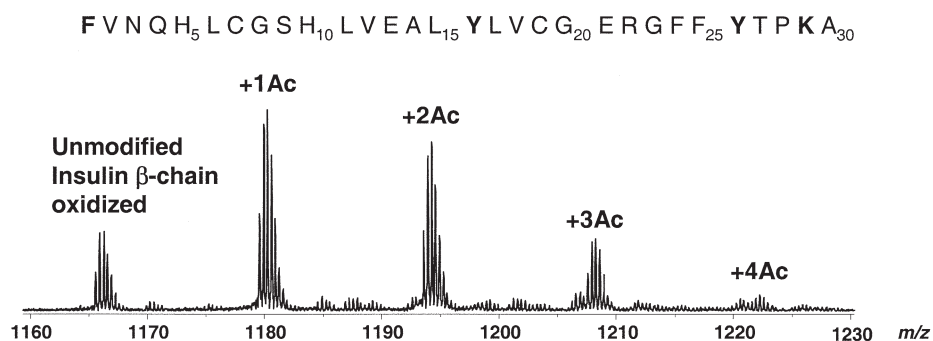$F V N Q H_5 L C G S H_{10} L V E A L_{15} Y L V C G_{20} E R G F F_{25} Y T P K A_{30}$

Figure 2. Mass spectrum of the reaction mixture of Insulin β-chain oxidized with NHSAc (1:3 ratio of peptide:NHSAc, pH 26> 8.4), the 3+ charge state is shown.

probe the reactivity of each type of residue, MS/MS spectra of the singly- and doubly-acetylated insulin products were measured, and regions around the y4, y6, and b15 ions that can be used to determine the modification status of the Phe 1, Tyr 27, and Lys 29 respectively are shown in Figure 3. From the studies with angiotensin I, it was shown that under all pHs studied, the primary amine at the amino terminus is more reactive than the tyrosine residues. This trend is mirrored when investigating oxidized insulin β-chain, as is shown in Figure 3. For example under acidic conditions, the N-terminus is ~80% modified as compared to Tyr 27, which is ~20% modified (Figure 3b).

This is determined through comparison of the ratio of y6:y6 + Ac (4:1) and b15:b15 + Ac (2:5). Under alkaline conditions (Figure 3d), virtually no tyrosine is modified (since y4 + Ac and y6 + Ac are approximately equal, therefore all contribution to y_6+Ac comes from the modified lysine), as compared to ~10% modification of the N-terminus. When comparing the reactivity of the Phe 1 and Lys 29, it is clear that the N-terminus is more reactive under acidic conditions. This is illustrated in Figure 3b and c where y4 + Ac is absent, or present at c.a. 10% relative abundance when the singly- and doubly-acetylated precursors from acidic conditions are subjected to tandem mass spectrometry, respectively.
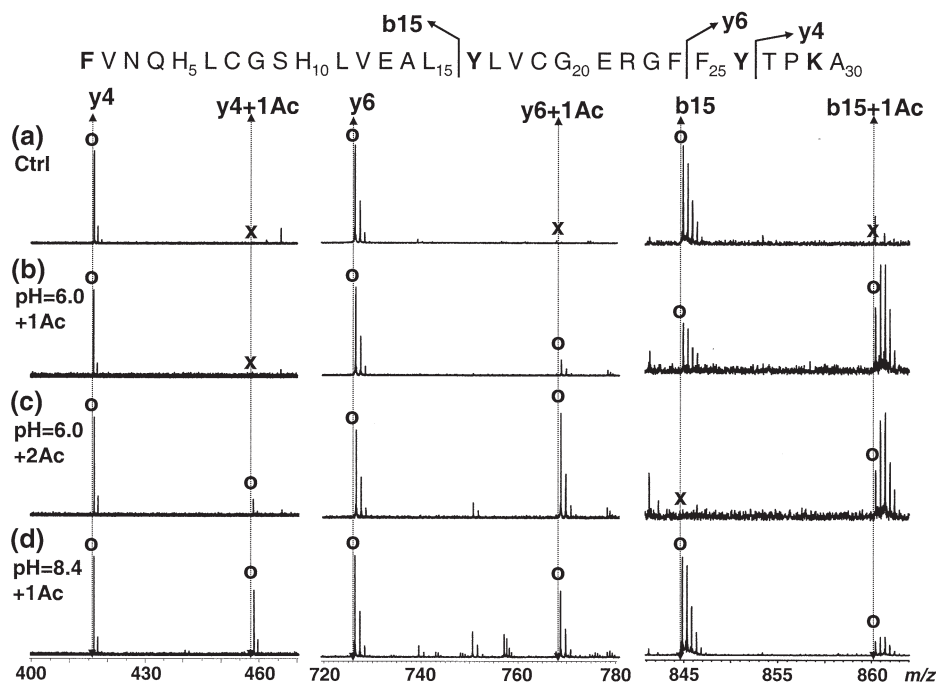

Figure 3. Expansions of MS/MS spectra of (**a**) unmodified oxidized insulin β-chain, (**b**) singly-acetylated oxidized insulin β-chain at pH =6.0, (**c**) doubly-acetylated oxidized insulin β-chain at pH = 6.0, and (**d**) singly-acetylated oxidized insulin β-chain at pH = 8.4. For each of the different reaction conditions, three spectral regions are shown that indicate the modification status of specific residues in the peptide. The y4 and y6 ions are singly-charged, while the b16 ion is doubly-charged. In all reactions the peptide to cross-linker ratio was 1:3. Note, an "**x**" indicates the absence of an ion, while an "**o**" indicates the presence of an ion.

On the other hand, under alkaline conditions it appears that Lys 29 is more reactive (Figure 3d), since the b15 + Ac ion is about 10% relative abundance as compared to the unmodified b15 ion, and the ratio of the y4 and y4 + Ac ions is approximately 6:4.

When comparing the reactivity of lysine 29 and tyrosine 27 with the succinimidyl esters, a clear pH dependence emerges. At pH = 6.0, tyrosine modification predominates in both the singly- and doubly-acetylated precursors (Figure 3b, c). This is determined by comparing the ion intensities of the y6 + Ac ion as compared to the y4 + Ac ion. In each case the y6 + Ac ion is in greater abundance than the y4 + Ac ion. However, at pH = 8.4, the y4 + Ac:y4 and the y6 + Ac:y6 intensity ratios are approximately equal (Figure 3d). Since acetylated y6 can come from modification of Lys 29 and Tyr 27 and an increase in the y6 + Ac:y6 ratio is not observed, the majority of the modified product must be coming from modification of Lys 29 at pH 8.4.

From the investigation of the succinimidyl ester reactivity with oxidized insulin β-chain and angiotensin I, it is clear that the original reactivity order proposed by Staros holds in the present experiments [12, 13]. Taken together, these results indicate the opportunity to generate different types of cross-links at differing pH. From the above data, it is likely that under acidic conditions chemical cross-links will be preferentially formed between the N-terminus and tyrosine or between tyrosine-tyrosine. However, under alkaline conditions either lysine-lysine or lysine-N-terminus cross-links will be favored. The observed differences in reactivity can be explained by the nucleophilicity and the pKa of the reactive residues. When neutral, the nucleophilicity of the N-terminus and lysine is ~100-fold greater than the nucleophilicity of tyrosine. However, by varying the pH of the reaction, the reactivity of the primary amine-containing residues can be metered, since it is the neutral form of the primary amines that reacts with the succinimindyl esters. Therefore lysine (pKa ~9.5) is virtually unreactive at acidic pH and quite reactive at basic pH; while the N-terminus (pKa ~7.6) is reactive at both of the pHs studied. Furthermore, the smaller abundance of nucleophilic amines at acidic pH allows the tyrosine reaction to effectively compete with the primary-amine reaction.

Reaction of Ethylene Glycol Bis [Succinimidylsuccinate] with Oxidized Insulin β-Chain at Alkaline and Acidic pH

To determine the feasibility of Tyr-Tyr cross-linking, the cross-linking reaction of Ethylene glycol bis [succinimidylsuccinate] (EGS) with oxidized insulin β-chain was explored at both acidic and alkaline conditions. In choosing a suitable cross-linker for a feasibility study, the distance of the cross-linkable residues must be considered. In the crystal structure of insulin, the larg-

est cross-linkable distance is between the phenolic oxygens of Tyr 16 and Tyr 26 and is ~17 Å [16]. However, in NMR studies of oxidized insulin β-chain, it is clear that the peptide is extremely conformationally flexible [17]. In order to study the reactivity of the succinimidyl ester cross-linker with oxidized insulin β-chain, and in particular to determine if a tyrosine-tyrosine cross-link was possible, EGS was chosen. EGS has a 13-atom spacer arm with ~16 Å cross-linking distance [15]. With the distance of the cross-linker and the conformational flexibility of the peptide chain, it is expected that a tyrosine-tyrosine cross-link will form if the intrinsic reactivity of the succinimidyl esters will allow this to occur. Upon reacting oxidized insulin β-chain with EGS, both the singly- and doubly-cross-linked products were observed, which is consistent with the four reactive residues found in the acetylation study. The b16 ion was used to monitor the formation of the Phe 1-Tyr 16 cross-link, and the y6 ion was used to monitor the formation of the Tyr 26-Lys 29 cross-links, (Figure 4). Upon subjecting the singly-modified precursor to collisional activation, it was apparent that the Tyr 26-Lys 29 cross-link was preferentially observed under alkaline conditions, as the abundance of the y6 + EGS product ion is greater in alkaline conditions (Figure 4e) than in acidic conditions (Figure 4c). Conversely, the Tyr 26-Lys 29 cross-link was almost non-existent at the acidic pH investigated (Figure 4c). However at acidic pH, the modified b16 ion is quite abundant owing to reactivity of the Tyr 16 and the N-terminus at acidic pH (Figure 4d). While the Phe 1-Tyr 16 cross-link still forms at alkaline pH, it does so in less abundance (Figure 4f).

While the mixed cross-links were of interest, it is also important to determine if tyrosine-tyrosine cross-links can be formed. The internal ions 3-27-H$_2$O and 3-27, contain *only* Tyr 16 and Tyr 26 and were used to determine formation of tyrosine-tyrosine cross-links (Figure 5). At acidic pH, both the 3-27-H$_2$O + EGS and the 3-27 + EGS internal product ions are observed, indicating that the two tyrosines are cross-linked (Figure 5b). However, under alkaline conditions, a small amount of the cross-linked product is also observed (Figure 5c). Thus as is expected from previous data, it is possible to cross-link tyrosines with succinimidyl ester cross-linkers, and the reaction of tyrosine is favored at acidic pH. Therefore, it is possible to utilize the succinimidyl ester cross-linkers as if they are both hetero- and homo-bifunctional reagents, cross-linking both primary amines and tyrosines.

It is important to consider, that in contrast to the labeling experiments, the intensities of the unmodified product ions in a cross-linking experiment cannot *always* be used to determine the relative amounts of each type of cross-link. This is due to the overlapping of the cross-linked amino acid residues involved in a given product ion. For example, at pH = 6.0 it could appear that 100% of the cross-link was formed between Phe 1-Tyr 16 since there is no unmodified b16 ion (Figure 4d). However, there is also a large amount of modified
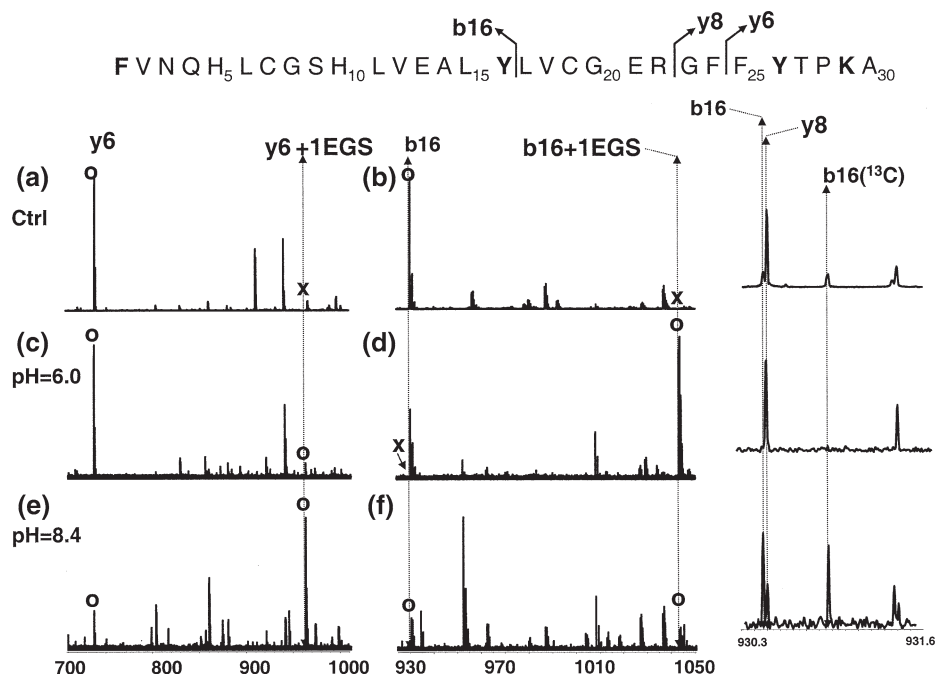
F V N QH_5 L C G SH_{10} L V E A L_{15} Y L V C G_{20} E R G F F_{25} Y T P K A_{30}

Figure 4. MS/MS of unmodified oxidized insulin β-chain in the region of the y6 ion (**a**), and the b16 ion (**b**). Product ion spectrum of the singly cross-linked oxidized insulin β-chain with EGS (1:3 ratio peptide:EGS) at pH = 6.0 (**c**, **d**) and pH = 8.4 (**e**, **f**). The inset shows the overlap of doubly-charged b16 ion (m/z 930.4415) and singly-charged y8 ion (m/z 930.4725). A concomitant decrease in the y8 ion is observed upon reaction with EGS as well. Note, an "**x**" indicates the absence of an ion, while an "**o**" indicates the presence of an ion.
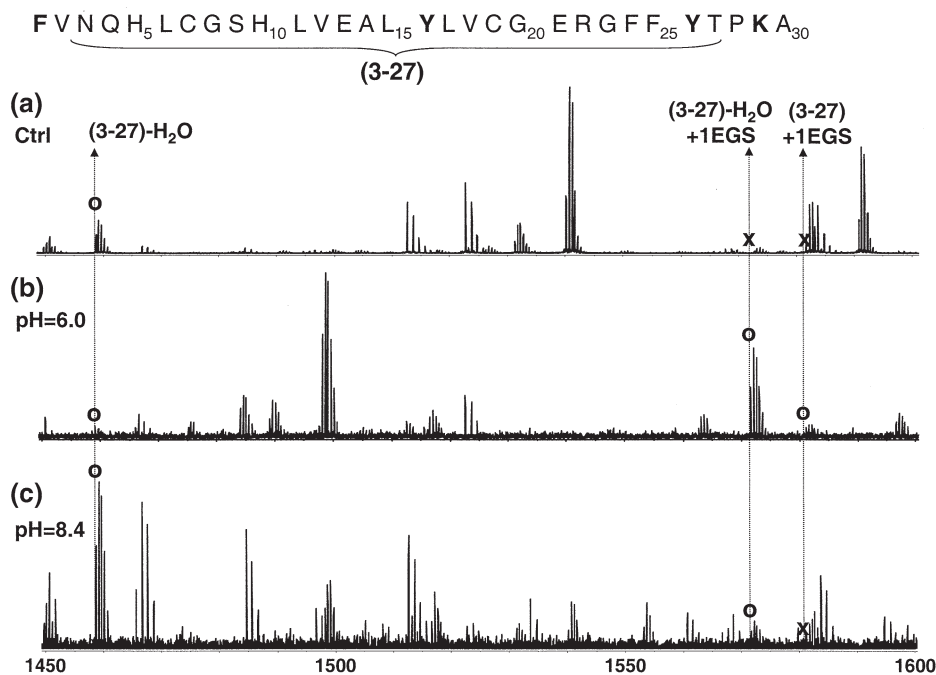
F V N QH_5 L C G SH_{10} L V E A L_{15} Y L V C G_{20} E R G F F_{25} Y T P K A_{30}

(3-27)

Figure 5. MS/MS of (**a**) unmodified oxidized insulin β-chain, singly-cross-linked oxidized insulin β-chain with EGS (1:3 ratio) at (**b**) pH = 6.0 and (**c**) pH = 8.4. Note, an "**x**" indicates the absence of an ion, while an "**o**" indicates the presence of an ion.

3-27-H_2O and very little unmodified 3-27-H_2O (Figure 5b). The formation of unmodified b16 is precluded due to the formation of the Phe 1-Tyr 16 cross-link and the Tyr 16-Tyr 26 cross-link.

Conclusions

Through the use of labeling and cross-linking reagents, the reactivity of succinimidyl ester cross-linkers was studied. Our labeling and cross-linking results are in agreement with the results of earlier studies on the reactivity of succinimidyl ester cross-linkers which showed that they react most readily and form stable products with primary amines and tyrosine. In addition, we found that the cross-links are sufficiently stable to be localized by MS/MS or proteolysis [12, 13]. Further, the specificity of these reagents was investigated at different pHs. Under acidic conditions (pH 6.0), the succinimidyl ester cross-linkers were found to react preferentially with the N-terminus and tyrosine, however under alkaline conditions (pH 8.4) they were found to react preferentially with the N-terminus and lysine. Using EGS and oxidized insulin β-chain, it was shown that not only could N-terminus/lysine-tyrosine cross-links be formed, but also tyrosine-tyrosine cross-links. This observation is of vital importance to chemical cross-linking studies, especially for the development of cross-linkers for MS3D studies of the structure of single proteins, as well as for the more established use of cross-linkers to study the assembly of multimeric protein complexes. In these studies, the correct assignment of reactive residues involved in a cross-link is crucial to correctly deriving distance constraints for use in developing a structural model for a single protein as well as for understanding the overall assembly of the multiprotein complex. Furthermore, utilizing the pH influence on the reaction, it is possible to use the succinimidyl ester cross-linkers in both hetero- and homobifunctional cross-linking studies, thus adding to the versatility of the reagent. When using changes in pH to affect the reactivity of different residues, the effect of the pH on protein structure must be considered. The solution conditions must be carefully chosen when analyzing protein conformations by numerous techniques including NMR (where NMR compatible solvents must be employed) or X-ray (where specific conditions are utilized to induce crystal formation). However, these changes are often subtle and would not be detected by the cross-linking method presented here, which is designed to determine the overall folding structure of the protein when combined with computational methods. Our future work will include the use of succinimidyl ester cross-linkers for lysine-lysine, lysine-tyrosine, and tyrosine-tyrosine cross-linking in MS3D studies.

Acknowledgments

References

1. Brunner, J. New Photolabeling and Cross-Linking Methods. *Annu. Rev. Biochem.* **1993,** *62,* 483–514.
2. Peters, K.; Richards, F. M. Chemical Cross-Linking: Reagents and Problems in Studies of Membrane Structure. *Annu. Rev. Biochem.* **1977,** *46,* 523–551.
3. Wold, F. *Methods in Enzymology, Vol. 25;* Academic Press: New York, **1972,** 623.
4. Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. High Throughput 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.
5. Schilling, B.; Row, R. H.; Gibson, B. W.; Guo, X.; Young, M. M. MS2Assign, Automated Assignment and Nomenclature of Tandem Mass Spectra of Chemically Cross-Linked Peptides. *J. Am. Soc. Mass Spectrom.* **2003,** *14,* 834–850.
6. Novak, P.; Young, M. M.; Schoeniger, J.; Kruppa, G. H. A Top-Down Approach to Protein Structure Studies Using Chemical Cross-Linking and Fourier Transform Mass Spectrometry. *Eur. J. Mass Spectrom.* **2003,** *9,* 623–631.
7. Kruppa, G. H.; Schoeniger, J.; Young, M. M. A Top Down Approach to Protein Structural Studies Using Chemical Cross-Linking and Fourier Transform Mass Spectrometry. *Rapid. Commun. Mass Spectrom.* **2003,** *17,* 155–162.
8. Collins, C. J.; Schilling, B.; Young, M.; Dollinger, G.; Guy, R. K. Isotopically Labeled Cross-Linking Reagents: Resolution of Mass Degeneracy in the Identification of Cross-Linked Peptides. *Bioorg. Med. Chem. Lett.* **2003,** *13,* 4023–4026.
9. 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.
10. Dihazi, G. H.; Sinz, A. Mapping Low-Resolution Three-Dimensional Protein Structures Using Chemical Cross-Linking and Fourier Transform Ion-Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2003,** *17,* 2005–2014.
11. Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. A Top-Down Method for the Determination of Residue-Specific Solvent Accessibility in Proteins. *J. Mass Spectrom.* **2004,** *39,* 322–328.
12. Anjaneyulu, P. S.; Staros, J. V. Reactions of N-Hydroxysulfosuccinimide Active Esters. *Int. J. Pept. Protein Res.* **1987,** *30,* 117–124.
13. Staros, J. V. Membrane-Impermeant Cross-Linking Reagents: Probes of the Structure and Dynamics of Membrane Proteins. *Acc. Chem. Res.* **1988,** *21,* 435–441.
14. Cuatrecasas, P.; Parikh, I. Adsorbents for Affinity Chromatography. Use of N-Hydroxysuccinimide Esters of Agarose. *Biochemistry* **1972,** *11,* 2291–2299.
15. Hermanson, G. T. *Bioconjugate Techniques, 1st ed.;* Academic Press: San Diego, **1996,** 139–140.
16. Gursky, O.; Badger, J.; Li, Y.; Caspar, D. L. Conformational Changes in Cubic Insulin Crystals in the pH Range 7–11. *Biophys. J.* **1992,** *63,* 1210–1220.
17. Hawkins, B.; Cross, K.; Craik, D. Solution Structure of the B-Chain of Insulin as Determined by 1H NMR Spectroscopy. Comparison with the Crystal Structure of the Insulin Hexamer and with the Solution Structure of the Insulin Monomer. *Int. J. Pept. Protein Res.* **1995,** *46,* 424–433.