A Novel Protein Cross-Linking Reaction in Stressed Neutral Protamine Hagedorn Formulations of Insulin

RONALD C. BEAVIS,* MICHAEL D. KNEIRMAN, DAVID SHARKNAS, MARK A. HEADY, BRUCE H. FRANK, AND MICHAEL R. DEFELIPPIS

Contribution from *Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, DC 3811, Indianapolis, Indiana 46285.*

Received June 22, 1998. Final revised manuscript received December 8, 1998.
Accepted for publication December 9, 1998.

Abstract □ The covalent insulin–protamine product molecules formed by heat stress in Neutral Protamine Hagedorn formulations of insulin and the insulin analogue [Lys^{B28},Pro^{B29}] were examined by mass spectrometry. The results demonstrated that the covalent cross-link between insulin and protamine was not caused by linkage through the protamine N-terminal amino group, as had been previously thought. Our results indicate that the linkage was formed between the side chain of a protamine arginine and a histidine in the insulin B chain, resulting in a net mass change of –5 Da, compared to the sum of the protamine and insulin molecular masses. A mechanism for this new type of covalent cross-linking reaction is proposed.

Introduction

The Neutral Protamine Hagedorn (NPH) formulation of insulin is the oldest and most successful example of a sustained release protein pharmaceutical product.[1] NPH[2] consists of insulin cocrystallized with a defined amount of the very basic protein protamine. The resulting microcrystalline suspension is administered to a patient parenterally. The crystals dissolve very slowly in the subcutaneous depot site, resulting in a protracted release of free insulin. The pharmacological properties of NPH have been thoroughly studied.[3]

The long history of the NPH formulation has led to careful examination of the products that form during long-term storage of this pharmaceutical preparation. Brange and co-workers[4−6] have examined these products in detail and determined that they fall into two general classes: modification of side chain moieties of insulin, and polymerization reactions that form both covalently cross-linked insulin dimers (CID) and covalently cross-linked insulin−protamine products (CIPP). The nature of the polymerization reactions that cause the insulin dimer products could not be directly determined by these studies. Brange et al.[4−6] speculate that the polymers were produced by the attack of an N-terminal amino group on a side chain amide group, resulting in the loss of ammonia and the formation of an isopeptide bond. This cross-linking reaction was referred to as either aminolysis or transamidation. It was further speculated that any of the six amide-containing residues of insulin could participate in these reactions leading to the observed chromatographic multiplicity of insulin dimer peaks. Formation of insulin−protamine polymer products was assumed to proceed by the same mechanism, i.e., transamidation between the N-terminus of protamine and any of the amide groups in insulin. The protamine used in pharmaceutical preparations of NPH does not have any amide moeities,[7] so there can be no reactions of this type between protamine and the N-termini of mature insulin.

Anderson and co-workers[8−10] have suggested a different reaction mechanism for the formation of covalent insulin dimers. Their mechanism requires the formation of a succinyl anhydride by the loss of ammonia from the C-terminal asparagine residue on the A-chain of insulin. Succinyl anhydride can then react with water to form $desamido_{A21}$-insulin. Alternatively, succinyl anhydride can react with a free amino group on the N-terminus of insulin's A or B chain to form a covalently linked dimer. This reaction can result in either a peptide or isopeptide bond linking the two chains together. Their work does not directly consider the formation of insulin−protamine products; however, it would be reasonable to assume that a highly reactive succinyl anhydride should cause a similar transacylation reaction with the free N-terminus of protamine.

Conventional reversed-phase HPLC methods result in many CID peaks and a rather ambiguous set of CIPP peaks. A size exclusion method has been recently developed to separate insulin from any covalent cross-linked products, in a solution formed by dissolving NPH insulin crystals in an acidic solution.[11] This method results in three resolved peaks: (1) monomeric insulin; (2) containing all CID species; and (3) containing all CIPP species. In this work, we have used this method to purify relatively large amounts of CIPP. The isolated material was characterized using a combination of mass spectrometry and standard protein chemistry techniques. Our results show that the CIPP peak consists of a mixture of many different insulin−protamine covalent dimers. We show that CIPP dimers are not linked by either a peptide or isopeptide bond. Instead, we propose that a condensation reaction occurs, involving any arginine group in protamine and the histidine groups on the B chain at position B5 or B10. This proposed reaction results in the loss of 5 Da, compared to the sum of the starting materials.

Experimental Section

Chemicals—NPH and NPL[12−14] (a similar preparation to NPH that uses the [Lys^{B28},Pro^{B29}] analogue of insulin) suspensions were obtained in house from Eli Lilly and Company (Indianapolis, IN) supplies. Chum salmon (*Oncorhynchus keta*) protamine was also obtained in house from Eli Lilly and Company (Indianapolis, IN). All chromatographic solvents were HPLC grade, except the water, which was obtained from a Milli-Q Plus water purification system (Millipore, Bedford, MA). Burdick & Jackson (Muskegon, MI) high purity grade water and acetonitrile were used for mass spectrometric sample preparation. The trifluoroacetic acid used was Sequanal grade (Pierce, Rockford, IL). *trans*-α-cyano-4-hydroxycinnamic acid[15] (Acros, Pittsburgh, PA) was dissolved as an ammonium salt, recrystallized with dilute hydrochloric acid and repeatedly washed with water prior to use as a MALDI matrix. Porcine trypsin was obtained from Promega (Madison, WI), and

10.1021/js9802603 CCC: $18.00
Published on Web 01/20/1999

endo-Glu-C was obtained from Sigma (St. Louis, MO). Ammonium bicarbonate (Fisher Scientific, Pittsburgh, PA) and dithiothreitol (Aldrich Chemical Co., Milwaukee, WI) were used in digesting the CIPP fractions.

Sample Preparation and Storage—One vial of NPH suspension contains the following ingredients: biosynthetic human insulin (3.5 mg/mL); protamine free base (0.27 mg/mL); dibasic sodium phosphate (3.78 mg/mL); glycerol (16 mg/mL); phenol (0.65 mg/mL); *m*-cresol (1.60 mg/mL); zinc oxide (to yield 0.025 mg/mL Zn^{2+}); water for injection. One vial of NPL suspension contains the following ingredients: biosynthetic [Lys^{B28},Pro^{B29}] human insulin analogue (3.5 mg/mL); protamine free base (0.30 mg/mL); dibasic sodium phosphate (3.78 mg/mL); glycerol (16 mg/mL); phenol (0.72 mg/mL); *m*-cresol (1.76 mg/mL); zinc oxide (as needed to yield 0.025/ml mg Zn^{2+}); water for injection.

CID and CIPP molecular species were generated in these vials of suspended crystals by heat stressing the sealed containers in normal atmospheric air for 3−6 months at 30−37 °C. These conditions normally promote mild oxidation of the tested protein, such as the formation of methionine sulfoxide, if methionine residues are present in the sequence.

Size Exclusion Chromatography—The method used has been previously described.[11] The chromatography system was composed of an AS-100T autosampler (BIO-RAD, Mountain View, CA), a Protein-Pak 125, 10 μm, 7.8 × 300 mm size exclusion column, a ConstaMetric 3500 pump (LDC Analytical, Riviera Beach, FL), and a variable wavelength UV detector (Applied Biosystems, Foster City, CA). The mobile phase was 65:20:15 5.74 mM l-arginine:acetonitrile:glacial acetic acid (v/v/v). The volume flow rate was 0.5 mL/min, the injection volume was 100 μL, and the column temperature was room ambient. The total run time for the method was approximately 40 min, and all peaks were collected manually. To improve CIPP recovery from dried fractions, the arginine was left out of the mobile phase for some sample collections. The absence of arginine from the mobile phase did not affect the method's ability to resolve the CIPP or CID peaks from monomeric insulin.

Enzymatic Digestion—Purified CIPP fractions were digested with trypsin by first vacuum evaporating 10 μL of a fraction in a 0.5 mL polypropylene Eppendorf tube, using a SpeedVac drying apparatus (Savant Instruments, Holbrook, NY), with no heating applied during drying. The resulting deposit was dissolved in 20 μL of 50 mM ammonium bicarbonate (pH = 8.0), and 1 μL of 0.1 g/L trypsin was added. The tube was incubated for 1 h at 37 °C. If the sample was to be reduced, sufficient dithiothreitol was added to bring its concentration to 10 mM and incubated for 20 min at 20 °C. For endo-Glu-C digestion, the same procedure was followed, except that 1 μL of 0.1 g/L endo-Glu-C was substituted for trypsin. All other relevant factors remained the same. Pharmaceutical grade biosynthetic human insulin was used as the control in all of the studies mentioned in this paper, with parallel digestions and analyses performed. The results of these controls were not included explicitly in the Results section, because they did not demonstrate any unexpected results.

Mass Spectrometry—ESI mass spectrometry was performed using an API-III triple quadrupole (PE-Sciex, Foster City, CA) equipped with a nanospray ion source. Data analysis of the ESI spectra was performed using *BioMultiView* (PE-Sciex, Foster City, CA), version 1.2. The MS/MS experiment described in the Results section was performed with *m*/*z* steps of 0.24 and 0.14 on the first and third quadrupoles, respectively. The collision gas used was argon.

MALDI mass spectrometry was performed using a Voyager linear time-of-flight mass spectrometer (PerSeptive Biosystems, Framingham, MA). The matrix material used was *trans*-α-cyano-4-hydroxycinnamic acid. Typical sample preparation[16] consisted of first preparing a stock solution of the matrix material—a saturated solution of the matrix in 2:1:0.001 water:acetonitrile: trifluroacetic acid (v/v/v), prepared at room temperature. This solution was mixed with the sample at a ratio of 10 parts matrix solution to 1 part sample solution (v/v), and a 1 μL aliquot of the mixture was dried on the sample stage of the mass spectrometer. Data analysis of the MALDI spectra was performed using *m*/*z* (ProteoMetrics, New York, NY), version 8.6. Molecular mass calculations were performed using *PAWS* (ProteoMetrics, New York, NY), version 8.3. All masses shown as labels in figures or in tables below are the true molecular mass for the peptide, calculated from the experimental *m*/*z* value by the following formula: $m = z \times (m/z - 1.0079)$. All peptide masses calculated correspond to the isotopically averaged, chemical mass of the peptide.

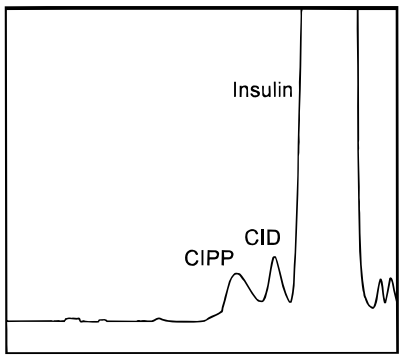

Figure 1—A typical size exclusion chromatogram of stressed NPH insulin, using the method described in the text. Free protamine was not recorded in this chromatogram because at the detector wavelength used (276 nm), protamine is transparent.

Table 1—The Molecular Masses of Chum Salmon Protamine

protamine variant	measured mass[a] (Da)	calculated mass[b] (Da)
1	4237	4236
2	4320	4319
3	4250	4250
4	4064	4064
5	4163	–
6	4418	–

[a] The measured masses were obtained by MALDI. [b] The calculated masses were obtained using the sequences and numbering system found in Hoffman,[7] except variants 5 and 6, which are reported here for the first time.

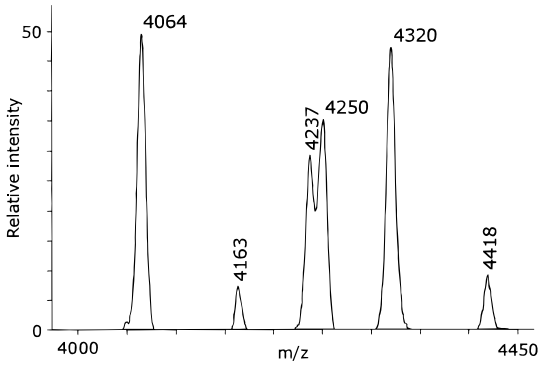

Figure 2—MALDI spectrum of protamine. See Table 1 for a comparison with known protamine sequences. The mass accuracy for the labeled peaks was ±1 Da.

Results

Figure 1 is a representative size exclusion chromatogram, demonstrating the resolution of the CID and CIPP species from insulin. Protamine, which is also present in the preparation, elutes just to the left of the CIPP peak. The detector wavelength used (276 nm) was chosen so that protamine itself is transparent, making the interpretation of the intensity of the CIPP peak unambiguous: the only absorption is the consequence of the presence of insulin[17] attached to one or more protamine molecules.

Even though protamine has considerable microheterogeneity[7] (see Table 1 and Figure 2), the size exclusion method produces only one peak corresponding to a dimer of itself with insulin. The presence of a single peak suggests that this product is the result of the addition of only one protamine molecule per insulin molecule. The mass spec-

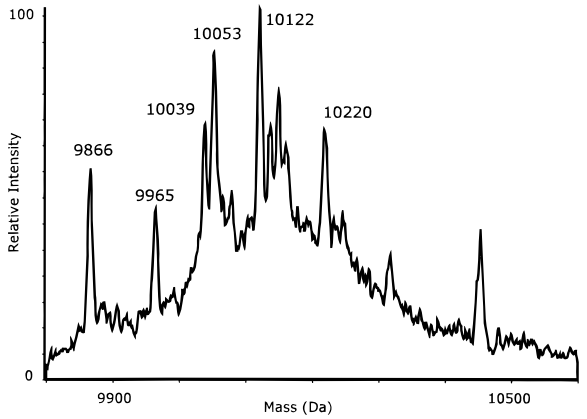

Figure 3—ESI spectrum of intact CIPP molecules. This spectrum was reconstructed from the multiply charged ion intensity pattern. The mass accuracy for the labeled peaks was ±1 Da.

Table 2—A Comparison of the Mass of the CIPP Dimer Observed by Electrospray Mass Spectrometry and the Molecular Mass Calculated by Simply Adding the Mass of Insulin to the Mass of a Protamine Variant (see Table 1)

observed dimer mass (Da)	assigned variant	protamine + insulin (Da)	Δ (Da)
10039	1	10044	–5
10122	2	10127	–5
10053	3	10058	–5
9866	4	9871	–5
9965	5	9971	–6
10220	6	10226	–6

trum in Figure 2 shows that the protamine used for the current study corresponded to known protamine sequences, although two new sequence variants (5 and 6) were discovered.

The reconstructed mass spectrum of the complete mixture of species found in the CIPP peak (see Figure 1) is shown in Figure 3. There are clearly many molecular species present in this mixture. Comparing the intensity distribution and molecular mass differences found in Figure 2 with those found in Figure 3, it was straightforward to assign which variant gave rise to a particular signal. For example, the lowest mass peak in Figure 3 (9866) is 99 Da lighter than the next peak (9965). This result corresponds to the same mass difference observed between the first two peaks in Figure 2. Table 2 shows the results of this correlation analysis. The first column is the observed mass of the CIPP species, column 2 is the assigned protamine variant, column 3 is the result of simply adding the mass of the protamine variant and insulin (5807.6 Da), and column 4 is the difference, $\Delta =$ (column 1) − (column 3). Examination of column 4 shows that the average mass loss in forming a CIPP dimer is ≈5 Da, regardless of which protamine variant is present in the CIPP dimer.

The spectrum shown in Figure 3 is the best electrospray spectrum obtained from CIPP material, but it still represents a very low intensity signal, relative to the baseline noise. MALDI produced much higher relative signal intensities for all CIPP samples, and it was the method of choice for analyzing these materials. In this case, however, the mass accuracy of the electrospray spectrum was sufficiently superior to the MALDI spectrum that it was used for the calculation shown in Table 2. The ionic species represented in Table 2 were those that were observed unambiguously in both the MALDI and electrospray spectra.

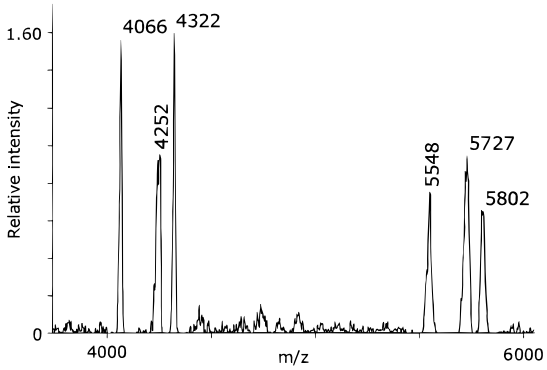

Figure 4—MALDI spectrum of CIPP, after treatment with endo-Glu-C and reduction with DTT. The mass accuracy for the labeled peaks was ±2 Da.

It is worth noting that even the purest CIPP fractions contained a significant amount of free protamine. The CIPP peak shown in Figure 1 elutes in the tail of the free protamine peak, which does not appear in the chromatographic trace because of the long wavelength chosen for the UV detector. All of the sample examined contained at least 30% free protamine, as estimated by MALDI-MS. Free protamine does not produce a useful signal in ESI-MS, so it could not be estimated by this method.

The nature of the cross-linking reaction was further investigated by performing a series of specific cleavages on the CIPP molecules and analyzing the results by mass spectrometry. The first of these cleavage reactions was to treat CIPP with the disulfide bond reducing reagent dithiothreitol and then cleave the resulting insulin A and B chains with the enzyme endo-Glu-C. This enzyme cleaves rapidly at the carboxy side of the peptide bond of glutamic acid residues, and more slowly at aspartic acid residues. Protamine does not contain any acidic residues, so it remains intact. The results of this cleavage are shown in Figure 4. The peaks in the mass range 4000−4400 Da correspond to unmodified protamine. The peaks in the mass range 5500−5900 Da correspond to protamine attached to the proteolytic fragment of insulin B chain, B[1−13]. These peaks also have $\Delta \approx -5$ Da. This result localizes the site of protamine attachment to somewhere in the first thirteen residues of the insulin B chain. Disulfide bond reduction was used to simplify the interpretation of the resulting molecular masses: the conditions used for this digestion reaction can promote disulfide shuffling.

One reasonable hypothesis regarding the attachment of protamine to insulin would be that the N-terminal amino group of protamine is involved in the cross-linking reaction. This hypothesis was tested by adding trypsin to the CIPP fraction, with the intention of cleaving all of the R−X (X ≠ P) bonds in protamine. The known protamine sequences[7] predicted that complete cleavage of all trypsin susceptible bonds would result in the dipeptide "PR" remaining attached to the insulin B chain. Dithiothreitol was added to the mixture after trypsin digestion to confirm that the modification was on the B chain.

Figure 5 shows the results of trypsin digestion followed by disulfide bond reduction. The signal $m/z = 2488$ and 3430 correspond to unmodified insulin B[1−22] and B[1−30], respectively. Table 3 assigns the signals for $m/z > 3430$ to possible tryptic fragments of the protamine sequences, along with the Δ value for each one of these assignments. Protamine produces many partially digested tryptic fragments with the same amino acid composition: the multiplicity per protamine molecule of possible fragments is given in the 4th column of Table 3, averaged over the known protamine sequences. The pattern of masses in the range $m/z = 2488-3327$ correspond to the same sequence

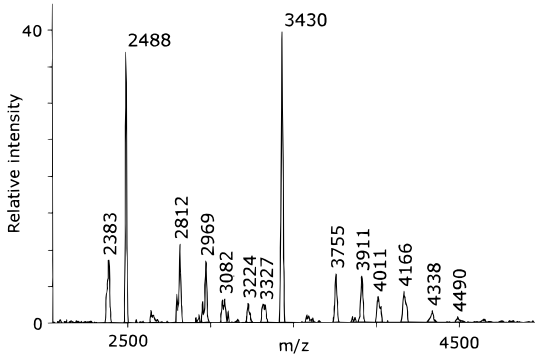

Figure 5—A portion of the MALDI mass spectrum of CIPP, following trypsin digestion and reduction with DTT. The mass accuracy for the labeled peaks was ±1 Da.

Table 3—Assignment of Tryptic Fragments of Protamine That Are Covalently Cross-Linked to the Human Insulin B Chain

measured mass (Da)	B chain + sequence	Δ (Da)	multiplicity[a]
3755	RR	–5	14
3911	RRR	–5	11
4011	PRRR	–5	2
4166	PRRRR; *or* RPRASR	–5	2
		–5	2
4338	RRRRGGR	–5	4

[a] Multiplicity indicates the number of different ways that a tryptic fragment with the same amino acid composition can be generated from the sequence of protamine.

and Δ assignments as in Table 3, except they are cross-linked to the tryptic fragment of insulin, B[1−22]. The signal at $m/z = 2383$ is the A chain of insulin.

The same types of experiments were performed using NPL formulations of the biosynthetic [Lys^{B28},Pro^{B29}] human insulin analogue. Identical results were obtained with the NPL formulation (data not shown).

Using the hypothetical reaction deduced in the section below, an MS/MS experiment was performed on CIPP material that had been subjected to a trypsin digestion, followed by a disulfide bond reduction and an endo-Glu-C digestion. The quadruply charged ion that would correspond to the B[1−13] cross-linked to the dipeptide RR less 5 Da ($m/z = 453.1$) was selected and MS/MS sequencing performed on that ion. The MS/MS data showed that the RR peptide was attached in the interval B[4−5]: the ion signals excluded all but these two residues as possible attachment sites. The interpretation of the signals from this MS/MS experiment was made possible by the deductions made below, rather than aiding in the formulation of these deductions, and therefore the data was not included in this study.

Discussion

The general strategy of using mass spectrometry on mixtures of peptides produced by enzymatic and chemical digests has been used to generate the main results in this study. The heterogeneity of the analytes and their physical properties make this approach the most viable method of gaining an understanding of the chemistry involved. Protamine and the protamine-insulin cross-linked species are very heterogeneous: there are at least 6 distinct protamine sequences and as many as 23 possible side chains that may participate in the proposed cross-linking reaction described below. Therefore, there are approximately 140 distinct molecular species that are being generically referred to as "CIPP". The samples that were analyzed in the course of this research were near the limits of detection for the mass spectrometric methods used, so any additional peak dilution caused by separating the individual components would compromise the quality of the data obtained from those peak fractions. Protamine-containing molecules do not separate well using reversed-phase separations because of the high solubility of protamine. Amino acid analysis was not useful for generating either positive or negative results because of the heterogeneity of the samples, their low concentrations, and the unusually high relative abundance of arginine in protamine-containing molecules, which leads to dynamic range problems in conventional amino acid analysis equipment.

Our results show that covalent CIPP species are present in stressed NPH insulin preparations and that they are very heterogeneous. The grossest level of heterogeneity results from the different sequences of protamine present in the initial preparation. The attachment site of protamine molecules to insulin is in the B chain of insulin. The results of endo-Glu-C digestion show that the protamine is linked to a residue somewhere within the first 13 amino acid residues of the insulin B chain. Tryptic digestion shows that the linkage can be made between B[1−13] and any arginine residue in protamine.

The value $\Delta = -5$ Da rules out the dimer formation reactions discussed by other authors[4−6,8−10] (see the Introduction). Transamidation reactions[4−6] result in the value $\Delta = -17$ Da. Succinyl anhydride-mediated reactions[8−10] result in $\Delta = -18$ Da. Cross-linking reactions resulting from the presence of low concentrations of aldehydes[4] would result in $\Delta \geq +12$ Da ($\Delta = +12$ is for the smallest aldehyde, formaldehyde). Protamine does not contain any cysteine residues, so disulfide cross-linking cannot occur.

The tryptic fragment assignments made in Table 3 rule out any dimer formation reaction involving the N-terminal amino group of protamine. If a free amino group was required to form the cross-link, then all of the sequences should begin with the N-terminal proline of the protamine sequence,[18] i.e., sequences beginning with the residues "PR ...". The observed sequences do not conform to this pattern. Instead they could come from any part of the protamine molecules, with fragments characteristic of both the N- and C-terminal region. Therefore, the protamine's N-terminal amino group cannot be involved in the dimer formation mechanism.

The fact that cross-linking to a single arginine residue is not observed (there is no signal at m/z = B chain + "R" − 5 Da) suggests that the cross-linking reaction involves the side chain of arginine in some way. The size of Δ suggests that the reaction involves the loss of 5 hydrogen atoms. This reaction would therefore not belong to the general class of condensation reactions typically found in protein chemistry. This different mechanism is probably brought about by the fact that this reaction is not a low temperature, solution phase reaction. Instead, this reaction occurs at elevated temperatures in the solid phase, i.e., within protamine-insulin crystals.

Inspection of the amino acid residues in the insulin B chain peptide B[1−13] can rule out a number of the residues for reaction with protamine. The sequence[16] is as follows:

FVNQHLCGSHLVE....

The side chains of phenylalanine, valine, leucine, and glycine residues are chemically inert under the conditions used to form CIPP. Therefore, the side chains of B1, B2, B6, B8, B11, and B12 do not participate in the reaction. The specificity of endo-Glu-C for cleaving at the carboxy side of glutamic acid residues implies that if B13 was modified there would be no cleavage at that site. The cleavage does occur; therefore B13 is not modified and does

not participate in the cross-linking reaction. Reactions known to occur with serine, asparagine, and glutamine residues all involve condensation reactions of their side chain amide groups, resulting in the loss of ammonia or water ($\Delta = -17$ or -18 Da). The results shown in Tables 2 and 3 rule out this type of reaction. Therefore residues B3, B4, and B9 do not participate in the formation of the CIPP observed in this study.

This simple process of elimination leaves the histidine residues, B5 and B10, and the cystine residue, B7-A7, as candidates to form of a cross-link with $\Delta = -5$ between a protamine guanidino group and insulin's B chain in a neutral pH, mildly oxidizing environment. There are no reactions that the authors are aware of (or can imagine) between an arginine guanidino group and a cystine residue that would result in this type of cross-linking under these conditions. Any reaction involving the oxidation of the sulfur atoms leading to a reactive product, such as cystine *S*-dioxide,[19] would result in $\Delta > 0$. A β-elimination reaction[20] involving that cystine could leave a reactive dehydroalanine at B7, but any reaction of that dehydroalanine with a guanidino group would result in $\Delta \leq -32$. These considerations eliminate the cystine residue and point to the involvement of at least one of the histidine residues.

Any dimer-forming reaction that occurs in the solid-state would be either assisted or inhibited by the geometry of insulin and protamine molecules in the NPH crystals. X-ray crystallographic examination of NPH crystals[21] shows that protamine does not occupy a single, well-oriented site in these crystals, due to its repetitive and heterogeneous structure. Electron density maps of these crystals show a region of protamine density at the interface between monomer 1 and the symmetry equivalent of monomer 3. It was suggested that the interactions between insulin and protamine molecules occurred via the guanidino groups of protamine's arginine residues and carboxylic acid containing residues in the insulin molecule. Contrary to this supposition, the actual structure shows that the most stable portions of the protamine electron density occur near the two histidine residues in the B chain, B5 and B10. The histidine residue B10 was also found to be responsible for the coordination of the zinc atoms in the assembled insulin hexamers.

Hydrogen bonding interactions between arginine and histidine side chains probably produce the localization observed in the X-ray crystal structure. Both histidine and arginine contain hydrogen bond donor and acceptor nitrogen atoms, because both contain singly bonded and doubly bonded nitrogens separated by one carbon atom. Therefore, there is the potential of forming two hydrogen bonds between these residues, which would lead to a significant stabilization of conformers that would allow the formation of these bonds in NPH crystals.

The most likely cross-linking reaction involves the imidazole group of histidine and the guanidino group of arginine, based on the combination of the known geometric association of insulin histidine and protamine arginine residues and the process of eliminating all of the other residues in the B[1−13] peptide. The mass lost during the reaction strongly suggests that two bonds are formed: a loss of four hydrogens leads to $\Delta = -4$, which is very close to the measured experimental value. However, the mechanism for this type of bond formation could not be probed directly in this system. Subsequently, the location of the cross-linking bridge was confirmed by LC/MS/MS (see Results) to be within the region B[4−5], supporting this hypothesis.

A number of reactions were considered for forming this cross-link, given the experimental constraints. Reactions involving the formation of a bridged tetrazine ring, and a

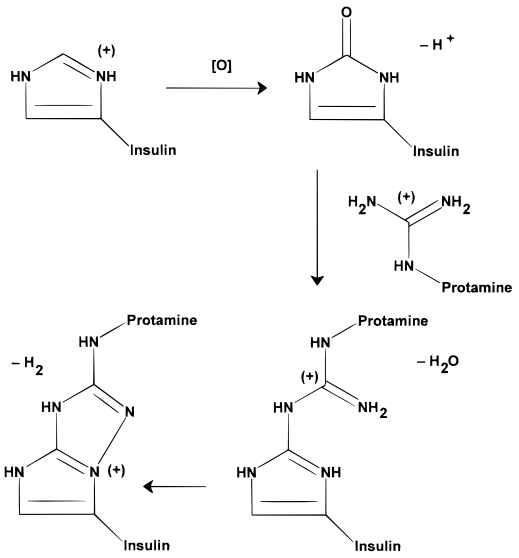

Figure 6—The proposed mechanism for the formation of the protamine-insulin cross-link, via an oxidation–reduction to form an imidazo[2,3-*b*]-1,2,4-triazolinium fused ring.[28]

number of other possible bridged rings were rejected because they would all result in violations of Bredt's rule.[22] One reasonable pathway to the formation of an appropriate cross-link is to postulate an initial oxidation of the histidine imidazole, as shown in Figure 6. This reaction, which forms imidazolone, is known to occur in histidine-containing peptides under oxidizing conditions.[23−25] This reaction has also been identified as a necessary step in the cross-linking of crystallins under oxidizing conditions.[26] The resulting imidazolone ring can be modified in several different ways, resulting in 2(3*H*)-, 2(5*H*)-, or 4(5*H*)-imidazolone.[27] Each of these exists as several tautomers,[26] by shifting a hydrogen to or from the keto group to form the tautomeric alcohol. The following discussion will center on the 2-imidazolone form because it has been observed in peptides.[23−25] Once the imidazolone derivative has been formed by oxidation, it is susceptible to a nucleophilic attack by one of the nitrogens in a protamine arginine side chain guanidino group. The final step in this tentative cross-linking mechanism is a further oxidation, as shown at the bottom of Figure 6.

This final step is necessary to explain the observed $\Delta = -5$ Da. The net result of the proposed reaction is the addition of oxygen (+16 Da), the loss of water (−18 Da), and the loss of H_2 (−2 Da), giving an overall $\Delta = -4$ Da. However, because the final product is positively charged, the additional proton that is normally assumed when making molecular mass calculations from experimentally determined *m/z* values is not necessary, resulting in the apparent loss of an additional proton, i.e., $\Delta = -5$ Da.

Conclusions

Whatever the exact mechanism, the results given in this paper clearly demonstrate that in the NPH formulation of insulin, the CIPP species formed are caused by a novel cross-linking reaction that occurs most rapidly under heat-stressed, mildly oxidizing conditions. Our observations do not agree with previous propositions as to the mechanism of CIPP formation, and they rule out most common condensation reactions that are known to occur between amino acid side chains. The site of the cross-linkage is proposed to be between any arginine side chain in protamine and the B[1−13] residues of insulin. We propose the reaction pathway indicated by Figure 6 as being consistent

with the experimental evidence. This mechanism may be the cause of dimer formation in other proteins, when they are stressed under similar conditions. Further work is ongoing to test this hypothesis.

References and Notes

1. Hagedorn, H. C.; Jensen, B. N.; Karup, N. B.; Wodstrup, I. Protamine Insulinate. *J. Am. Med. Assoc.* **1936**, *106*, 177−180.
2. The following abbreviations are used: NPH, neutral protamine Hagedorn; NPL, neutral protein LisPro; CID, covalent insulin dimer; CIPP, covalent insulin−protamine product; MALDI, matrix-assisted laser desorption/ionization; ESI, electrospray ionization; HPLC, high performance liquid chromatography. All molecular masses are reported in Daltons (Da). Mass-to-charge ratios (*m*/*z*) are dimensionless.
3. Felig, P. Landmark perspective: Protamine insulin. Hagedorn's pioneering contribution to drug delivery in the management of diabetes. *J. Am. Med. Assoc.* **1984**, *251*, 393−396.
4. Brange, J.; Havelund, S.; Hougaard, P. Chemical Stability of Insulin. 2. Formation of Higher Molecular Weight Transformation Products During Storage of Pharmaceutical Preparations. *Pharma. Res.* **1992**, *9*, 727−734.
5. Brange, J. Chemical stability of insulin. 4. Mechanisms and kinetics of chemical transformation in pharmaceutical formulation. *Acta Pharm. Nord.* **1992**, *4*, 209−222.
6. Brange, J.; Hallund, O.; Sørensen, E. Chemical stability of insulin. 5. Isolation, characterization and identification of insulin transformation products. *Acta Pharm. Nord.* **1992**, *4*, 223−232.
7. Hoffman, J. A.; Chance, R. E.; Johnson, M. G. Purification and Analysis of the Major Components of Chum Salmon Protamine Contained in Insulin Formulations Using High Performance Liquid Chromatography. *Protein Expression Purif.* **1990**, *1*, 127−133.
8. Darrington, R. T.; Anderson, B. D. Evidence for a Common Intermediate in Insulin Deamidation and Covalent Dimer Formation: Effects of pH and Analine Trapping in Dilute Acidic Solutions. *J. Pharm. Sci.* **1995**, *84*, 275−282.
9. Darrington, R. T.; Anderson, B. D. Effects of Insulin Concentration and Self-Association on the Partitioning of Its A-21 Cyclic Anhydride Intermediate to Desamido Insulin and Covalent Dimer. *Pharm. Res.* **1995**, *12*, 1077−1084.
10. Strickley, R. B.; Anderson, B. D. Solid-State Stability of Human Insulin II. Effect of Water on Reactive Intermediate Partitioning in Lyophiles from pH 2−5 Solutions: Stabilization against Covalent Dimer Formation. *J. Pharm. Sci.* **1997**, *86*, 645−653.
11. "Insulin Preparations, Injectable." PA/PH/Exp. 6I/T (96) 4, ANP. Test: Impurities with Molecular masses Greater Than That of Insulin. *Pharmeuropa* **1996**, *8(3)*.
12. DiMarchi, R. D.; Mayer, J. D.; Fan, L.; Brems, D. N.; Frank, B. H.; Green, L. K.; Hoffman, J. A.; Howey, D. C.; Long, H. B.; Shaw, W. N.; Shields, J. E.; Slieker, L. E.; Su, K. S. E.; Sundel, K. L.; Chance, R. E. Synthesis of a fast-acting insulin based on structural homology with insulin-like growth factor I. In *Peptides: Chemistry and Biology*; Smith, J. A., Rivier, J. E., Eds.; ESCOM: Leiden, 1992; pp 26−28.
13. Long, H. B.; Baker, J. C.; Belagaje, R. M.; DiMarchi, R. D.; Frank, B. H.; Green, L. K.; Hoffman, J. A.; Muth, W. L.; Pekar, A. H.; Reams, S. G.; Show, W. N.; Shields, J. E.; Slieker, L. J.; Su, K. S. E.; Sundel, K. L.; Chance, R. E. A−C−B human proinsulin: a novel insulin agonist and intermediate in the synthesis of human insulin. In *Peptides: Chemistry and Biology*; Smith, J. A.; Rivier, J. E., Eds.; ESCOM: Leiden, 1992; pp 88−90.
14. DeFelippis, M. R.; Bakaysa, D. L.; Bell, M. A.; Heady, Li, S.; Youngman, K. M.; Radziuk, J.; Frank, B. J. Preparation and Characterization of a Cocrystalline Suspension of [Lys^{B28},-Pro^{B29}]-Human Insulin Analogue. *J. Pharm. Sci.* **1998**, *87*, 170−176.
15. Beavis, R. C.; Chaudhary T.; Chait, B. T. α-Cyano-4-hydroxycinnamic acid as a matrix for matrix-assisted laser desorption mass spectrometry**.** *Org. Mass Spectrom.* **1992**, *27*, 156.
16. Beavis, R. C.; Chait, B. T. Matrix-assisted laser desorption ionization mass-spectrometry of proteins**.** *Methods in Enzymology*; Karger, B. L.; Hancock, W. S., Eds.; 1996; Vol. 270A, pp 519−551.
17. Sures, I.; Goeddel, D. V.; Gray, A.; Ullrich, A. Nucleotide sequence of human preproinsulin complementary DNA**.** *Science* **1980**, *208*, 57−59.
18. All 14 of the *Oncorhynchus sp.* protamine sequences currently available in GenBank (http://www.ncbi.nlm.nih.gov/Web/Genbank/index.html) begin with the sequence "PR...".
19. *The Merck Index*, 11th ed.; Budavari, S., Ed.; Merck & Co.: Rahway, NJ, **1989**; p 437.
20. Kosen, P. A. Disulfide bonds in proteins. In *Stability of Protein Pharmaceuticals, Part A, Chemical and Physical Pathways of Protein Degradation: Pharmaceutical Biotechnology;* Ahern, T. J., Manning, M. C., Eds.; Plenum Publishing: New York, 1992; Vol. 2, pp 31−67.
21. Balschmidt, P.; Benned Hansen, F.; Dodson, E. J.; Dodson, G. G.; Korber, F. Structure of Porcine Insulin Cocrystallized with Clupeine Z. *Acta Crystallogr.* **1991**, *B47*, 975−986.
22. Wiseman, J. R.; Pletcher, W. A. Bredt's rule. III. Synthesis and chemistry of bicyclo[3.3.1]non-1-ene**.** *J. Am. Chem. Soc.* **1970**, *92*, 956−962.
23. Uchida, K.; Kawakishi, S. Selective oxidation of imidazole ring in histidine residues by the ascorbic acid-copper ion system. *Jpn. Biochem. Biophys. Res. Commun.* **1986**, *138*, 659−65.
24. Lewisch, S. A.; Levine, R. L. Determination of 2-oxohistidine by amino acid analysis. *Anal. Biochem.* **1995**, *231*, 440−446.
25. Uchida, K.; Kawakishi, S. 2-Oxohistidine as a novel biological marker for oxidatively modified proteins. *FEBS Lett.* **1993**, *332*, 208−210.
26. Guptasarma, P.; Balasubramanian, D.;Matsugo, S.; Saito, I. Hydroxyl radical mediated damage to proteins, with special reference to the crystallins. *Biochemistry* **1992**, *31*, 4296−4303.
27. Hofman, K. *Imidazole and Its Derivatives: Part I*; Interscience Publishers: New York, 1953; pp 60−63.
28. Preston, P. N. *Condensed Imidazoles: 5−5 Ring Systems*; Interscience Publishers: New York, 1986; pp 330−345.

Acknowledgments

The authors are indebted to Al Pekar for his preparation of one of the samples used in this paper and to Christine O'Keefe for Edman sequencing work. They also thank Drs. Steve Baertschi, John Towns, and Ralph Riggin for their critical reviews and discussion of this manuscript.

JS9802603