Hierarchy of N-acylation Sites in Human Insulin Studied by RP-HPLC and Mass Spectrometry

[1]Hina Zain[Δ], [1]Munir Ahmad[Δ], [1] Qurra-tul-Ann Afza Gardner and [1,2] Muhammad Akhtar*
[1]*School of Biological Sciences, University of the Punjab, New Campus, Lahore 54590, Pakistan.*
[2]*Centre for Biological Sciences, University of Southampton, SO17 1BJ, UK.*
ma3@soton.ac.uk*

(Received on 26th August 2015, accepted in revised form 30th September 2015)

Summary: Acylation of human insulin was studied using, C_2 and C_{12} transfer reagents, N-succinimidyl acetate and N-succinimidyl dodecanoate (lauryl) respectively. In general, the acylation reaction was more specific at high pH values (12.00) for the modification of the ε-amino group of K^{29B}. At pH around 8.5, multiple modifications occurred at other sites, in particular the N-terminal glycine of A-chain (G^{1A}). WithN-succinimidyl dodecanoate, about 88 % of the acylation was established to be at the ε-amino of K^{29B}, as shown using tryptic digestion and identification of the products by mass spectrometry. While with N-succinimidyl acetate, singly acetylated product with acetyl group at the ε-amino of K^{29B}, as well as doubly modified species containing acetyl group in the ε-amino group and the amino-terminal of G^{1A}were produced in equal amounts. The high regiospecificity of modification at the ε-amino group, by the C_{12} reagent, is attributed to steric factors. Moreover, the above investigation also highlights the purification regime of singly and doubly acetylated products on RP-HPLC using a shallow gradient, though the difference between two species, in the case of acetyl group, is of 43 Da. The overall results allow us to establish the hierarchy of accessibility of the three amino group; ε-amino group K^{29B} is the least hindered and accessible to C_2 as well as C_{12} reagents, then is the amino group of G^{1A} which is accessible to the C_2 but not the C_{12} reagent, finally that of F^{1B} is most hindered and accessible to neither.

Key Words: Mass spectrometry, Regiospecificity of acylation, Mono-acetyl insulin, Mono-dodecanoyl (lauryl) insulin.

Abbreviations:

G^{1A}, K^{29B}: Position of residue in A or B chain; MALDI-TOF: Matrix assisted laser desorption ionization time of flight, RP-HPLC: reverse phase high performance liquid chromatography; LC-MS: liquid chromatography mass spectrometry

Introduction

One of the approaches to increase the biological half-life of peptide drugs is to produce conjugates which either protect the drug against *in vivo* proteolysis or delay its clearance from the kidney [1-3]. Another strategy is the attachment of hydrophobic appendage which allows the drug to associate with albumin, from where it dissociates slowly [4, 5]. Examples of the latter are the acylation of insulin at the ε-amino group of lysine at position 29 in the B-chain or another lysine residue introduced in the A-chain [6]. Apart from the ε-amino group of Lys^{29B}, native insulin contains two other sites for potential acylation; these are the N-terminal amino acid residues of A and B chains (Gly^{1A} and Phe^{1B} respectively).

In this paper, we investigate the regiospecificity of acylation reaction using reagents of two different chain lengths and also explore the methods for the preparative separation of these hydrophobic derivatives.

Experimental

Materials

Esters of N-hydroxysuccinimde were prepared as described previously [8], human insulin and other chemicals were from Sigma.

Mono- and di-acetyl insulin

To 5 mg of human insulin in 2.70 ml of 20 mM sodium borate buffer (pH 12.0) containing 25% acetonitrile was added 0.625 ml of a solution of N-succinimidyl acetate (40 mM stock solution in 100% acetonitrile) and the mixture, in which the final concentration of the reagent was 3 mM and of acetonitrile 30.5%, was incubated at 25 °C for 10 min. 250 µl of 10% trifluoroacetic acid (TFA) was then added to quench the reaction. The reaction mixture was centrifuged at 10, 000*g* for 3 min and filtered through a 0.45µm filter assembly

*To whom all correspondence should be addressed; ΔThese authors have made equal contribution to the work.

(Minisartorius). A 50 µl aliquot of the latter mixture was desalted using G25 spin column (GE healthcare) and subjected to analysis by MALDI mass spectrometry. The remaining sample was then purified by reverse phase HPLC (SYKAM, Germany GmbH) using Biobasic C18analytical column (250 x 4.6 mm) with a gradient of 100% solvent A (0.1% TFA in water) for 3 min, and 30 – 50% of solvent B (99.9% acetonitrile in 0.1% TFA) for 43 min (0.5% ramp per min). HPLC gave three peaks at 38-38.5%, 38.5-39% and 39.3-40.3% of solvent B which on MALDI analysis gave $[M+1H]^+$ ions corresponding to mono-acetyl insulin (5850.6; theoretical, 5850.6), di-acetyl insulin (5892.6; theoretical 5892.6), and a mixture of di-acetyl and tri-acetyl insulin (5934.7; theoretical 5934.6). UV quantification showed that the yield of the purified mono-acetyl and di-acetyl insulin was about 30% each (based on the molar extinction co-efficient of insulin giving an OD_{280} of 1 for a 1 mg/ml solution; http://www.us.expasy.org). The HPLC fractions were lyophilized and stored at 4 ^oC.

Dodecanoyl (lauryl) insulin

For the preparation of dodecanoyl insulin, the modification was performed using 3 mM N-succinimidyl dodecanoate as above except that the incubation was at 25 ^oC for 30 min. Following the completion of the reaction the products were separated by RP-HPLC, using acetonitrile as a mobile phase as explained above but using a C4 column, two major peaks were obtained, peak 1 eluting at 50– 52% and peak 2 eluting at 57-58% of acetonitrile in 0.1% TFA. Peak 1 gave the $[M+1H]^+$ for mono-dodecanoyl insulin (5990.6; theoretical 5990.6) and peak 2 was identified as di-dodecanoyl insulin (6173.5; theoretical 6173.6). In ethanol gradient system, solvent A was 30% aqueous ethyl alcohol/0.1% TFA and solvent B was 95% ethyl alcohol/ 0.1% TFA, two major peaks were obtained, eluting at 66.6-68.1% of ethanol/ 0.1% TFA which corresponded to mono-dodecanoyl insulin and 70.6-71.9% of ethanol/ 0.1% TFA was identified as di-dodecanoyl insulin. In both cases, the isolated yield of mono-dodecanoyl insulin (peak 1) was 40% and di-dodecanoyl insulin (peak 2) yield was 6%.

MALDI mass spectrometric analysis of acylated insulin derivatives

Following acylation, 2 – 3 µl of desalted crude mixture (as mentioned above) and HPLC fractions (containing 0.6 – 1µg/µl of protein) were analyzed on Autoflex^{TM} TOF TOF Smartbeam 200 (Bruker, Germany GmbH) as detailed in [9].

Thiolytic cleavage analysis of acylated insulin

A portion of freeze-dried acylated protein was dissolved in 0.1% TFA and an aliquot containing 10 µg/15 µl was taken and subjected to reduction by the addition of 1.5 µl of 100 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEPH), prepared in distilled water. The reaction mixture was kept at 55 °C for 45 min and MALDI analyses were performed by mixing 1.5 µl of reaction mixture with 19 µl of sinapinic acid solution. The data were analyzed in linear detection mode, both under negative and positive mode.

It is to be noted that with both the modified proteins, the A and B chains were seen in negative and positive mode whereas modified A-chain with the dodecanoate moiety was only visible in negative mode. Hence, the data analyzed on MALDI-TOF for thiolytic cleavage analysis of acylated insulin derivatives are shown for negative mode.

MALDI results for thiolytic cleavage of mono-acetyl insulin showed single addition of acetyl group (+42) on B-chain with a $[M - 1H]^{-1}$ ion at 3471.1 (theoretical, 3471.5) whereas di-acetyl insulin exhibited acetylation at both the A and B chains with ionic peaks at 2424.5 and 3471.1 respectively (theoretical, 2424.7 and 3471.5). Similarly, thiolytic cleavage of mono-dodecanoyl insulin showed addition of a single dodecanoyl moiety (+182) on B-chain with a $[M - 1H]^{-1}$ ionic peak of 3611.16 (theoretical, 3611.5).While di-dodecanoyl insulin displayed modification of both A and B chains with the negative ionic peaks at 2564.5 and 3611.3 respectively (theoretical, 2564.7 and 3611.5).

Regiospecificity and Fingerprinting of Acylated Insulin by Proteomics Analysis

A portion of HPLC purified, freeze dried fraction of mono-acetylated and mono-dodecanoyl insulin was dissolved in 50 mM ammonium bicarbonate pH 8.4, an aliquot containing 50 µg of each derivative was taken, to which 2.5 µl of 0.1 M dithiothreitol (DTT) in 50 mM ammonium bicarbonate pH 8.4 (final concentration of DTT was 5 mM) was added and volume of reaction mixtures were made to 50 µl with 50 mM ammonium bicarbonate pH 8.4. The reaction mixtures were incubated at 55 ^oC for half an hour. The reduced protein solutions were brought to room temperature and acetonitrile was added up to 10%, followed by an addition of 10 µl of 0.1 M iodoacetamide, prepared in 50 mM ammonium bicarbonate pH 8.4 (final

concentration of iodoacetamide was 17 mM), and incubated at 37 °C for half an hour in the dark. Following alkylation, protein solutions were subjected to tryptic digestion, by the addition of 2.5 µg of sequencing grade porcine trypsin from Promega®. The reaction mixtures were incubated at 37◦C for 3 - 4 h and quenched by the addition of TFA (final concentration 1%; pH 2.0).

The on-line LC MS analysis of the reduced, alkylated, tryptic digested mixtures were performed as described previously [10], on 6224 TOF LC/MS (Agilent Technologies, USA), except that the TIC (total ion current) chromatogram was analyzed using Agilent MassHunter Qualitative AnalysisBioConfirm software (version B.02.00) and processed using isotopic deconvolution algorithm.

In case of mono-acetylated insulin, following reduction, alkylation and tryptic digestion monoisotopic masses of 2600.27 Da, 2610.09 Da, and 1001.49 Da were obtained (theoretical, 2600.26, 2610.08 and 1001.47). The fragment of mass 1001.49 corresponded to octapeptide of insulin B-chain from residues G^{23B} to T^{30B} (see below) containing a signature acetylated moiety attached on ε-amino group of lysine at position 29. The counter fragments of 2600.27 Da corresponded to a reduced alkylated, des-octa B-chain (F^{1B}-R^{22B}) and of 2610.09 Da to A-chain without modification. In case of mono-dodecanoyl insulin, apart from reduced alkylated A-chain and des-octa B chain, the signature octapeptide fragment (residues G^{23B} to T^{30B}) of monoisotopic mass of 1141.67 Da (theoretical, 1141.47) was observed which showed the dodecanoyl moiety to be attached to ε-amino group of lysine at position 29.

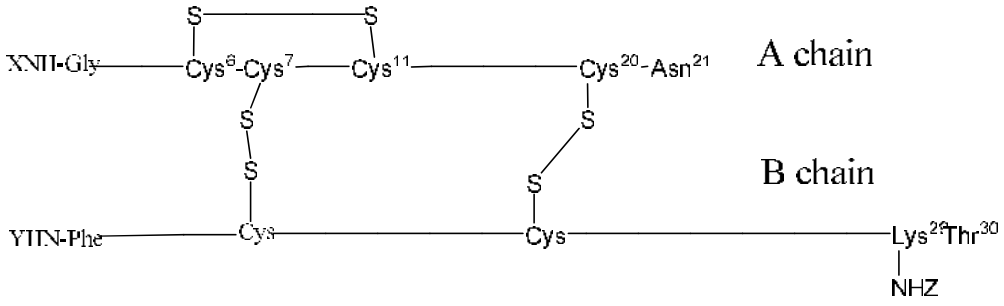
The origin of the tryptic peptides above may be deduced from the sequences of A- and B-chains (below): tryptic sites in the B-chain are present at R^{22B} and K^{29B} (shown by arrow),however, modification of the ε-amino of K^{29B} prevents the cleavage of the K^{29B}-T^{30B} bond, by trypsin, which would occur in the native B-chain.

A-chain H_2N-GIVEQCCTSICSLYQLENYCN-COOH

B-chain H_2N-F^{1}VNQHLCGSHLVEALYLVCG $ER^{22}\uparrow G^{23}$FFYTPK$^{29}\uparrow T^{30}$-COOH

Results and discussion

The reaction of N-succinimidyl acetate with insulin was studied using different ratios of the reagent and protein and at various pH values; the most favorable condition was the use of 10: 1 ratio of the reagent: insulin, with incubation time of 10 min in 20 mM sodium borate buffer, pH 12.0. The MALDI analysis of the crude reaction mixture showed the disappearance of all the insulin and the formation of mono- and di-acetyl insulin in about equal amounts, while tri-acetyl insulin was present as a minor product (Fig.1a). Using RP-HPLC with a shallow gradient of 0.1% TFA in acetonitrile and 0.1% TFA in water, and a ramp of 0.5%/min, mono-acetyl insulin (**2**, Scheme 1) was narrowly separated from the di-acetyl derivative, **(3)**, (Fig. 2a,peaks 1 and 2 respectively),while tri-acetyl insulin eluted in peak 3. These peaks could be separately collected and characterized by MALDI analysis, giving the expected masses for the three species, except that tri-acetyl insulin was contaminated with the di-acetyl derivative (Fig 1, traces b-d).



Scheme-1: Partial structure of human insulin **(1, X=H, Y=H, Z=H)** and its acylated products.

(2), X=H, Y=H, Z=CH_3-CO-
(3), Y=H, X=CH_3 –CO-, Z=CH_3 –CO-
(4), X=H,Y=H, Z=CH_3-$(CH_2)_{10}$-CO-
(5), Y=H, Z=CH_3-$(CH_2)_{10}$-CO-, X=CH_3-$(CH_2)_{10}$-CO-

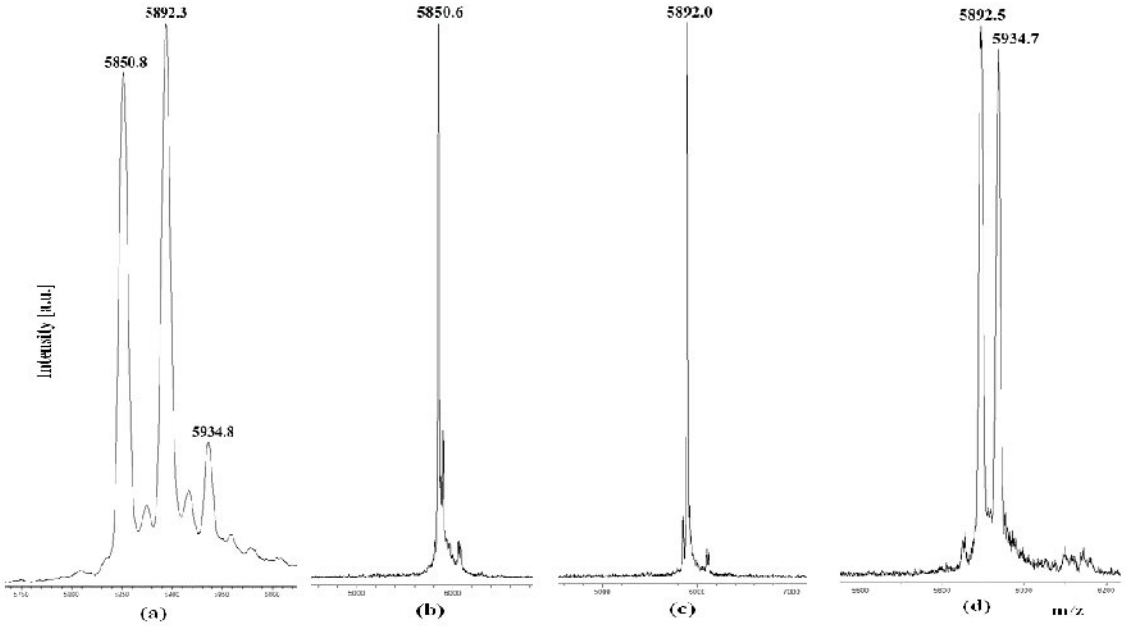


Fig. 1: Identification of acetylated products by MALDI. (**a**), $[M + 1H]^{+1}$ ions in crude acetylated insulin at 5850.8 (mono-acetyl insulin), 5892.3 (di-acetyl insulin), and 5894.8 (tri-acetyl insulin); **(b),** at 5850.6 (mono-acetyl insulin, peak 1, Fig. 2**a**); **(c)** at 5892.0 (di-acetyl insulin, peak 2, Fig.2**a**) and **(d)** at 5892.5 and 5934.7 (di-and tri-acetyl insulin, peak 3, Fig.2**a**).

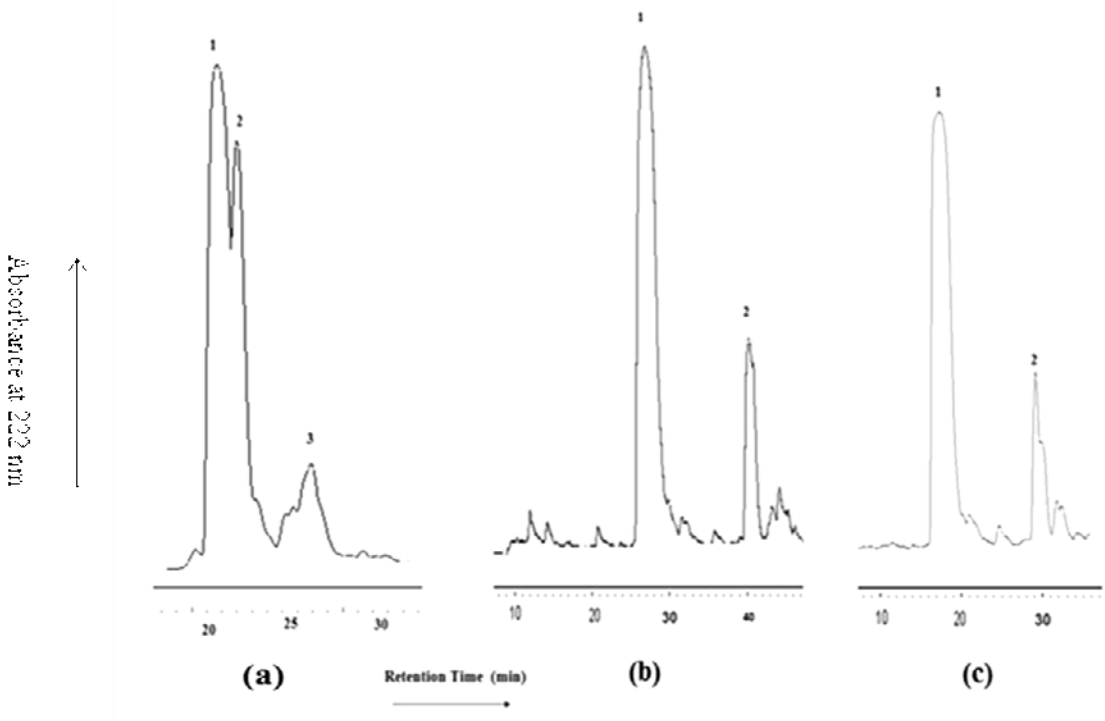


Fig. 2: RP-HPLC of standard insulin modified with N-succinimidyl acetate **(a)** and N-succinimidyl dodecanoate (**b** and **c**; with acetonitrile and ethanol gradient respectively).The analysis was performed on C18 column in **(a**) and on C4 column in **(b)** and **(c)** as described in the Methods section.

Next, the above protocol was extended to acylation using reagent with a C_{12} chain length, N-succinimidyl dodecanoate. Here, preliminary experiments showed that the reaction at pH 8.0 in 20 mM sodium phosphate buffer, gave a rather complex profile of multiple modifications which could be improved, in favor of predominant mono-acylation, by increasing the pH of the acylation medium. The optimal reaction conditions were found to be the use of 10: 1molar ratio of the reagent: insulin at pH 12.0 in 20 mM sodium borate buffer. The HPLC profile (Fig.2b) together with the MALDI spectrum of a typical experiment in Fig.3 shows the presence of mono- as well as di-dodecanoyl species**(4, 5)**, with the predominance of the former. The two could be separated by HPLC using the traditional acetonitrile-water gradient (Fig.2b), however, with this system the column tended to get blocked, after a few runs, presumably due to the presence of hydrophobic impurities arising from higher order modification by the C_{12} reagent. A safer option was the use of a gradient of 30% ethanol in water and 95% ethanol, which also gave good separation of the mono- and di-dodecanoyl insulin (Fig. 2c).

Thiolytic cleavage of these derivatives, using tris (2-carboxyethyl)phosphine hydrochloride led to the separation of the two chains and showed that the mono-acetyl insulin**(2)** in peak 1 (Fig. 2a) contained the acetyl moiety only in the B-chain (Fig. 4a), which was located at the ε-amino group of K^{29B}. While the di-acetyl-insulin **(3)** in peak 2 (Fig. 2a) following separation of the two chains was acetylated in both the chains (Fig. 4b); the presence of unmodified A-chain at 2382.8, in the sample, suggests contamination with the singly modified insulin.

The thiolytic cleavage of mono-dodecanoyl insulin**(4)** showed that the modification was on the B-chain (Fig. 4c) and its tryptic digest analysis, following thiolysis, established that the residue modified by the reagent was present in the octa peptide fragment constituting residues G^{23B} to T^{30B} in the B-chain of insulin (see Method section). Since the only amino group in this part of insulin is the amino group of K^{29B} this must have been acetylated.

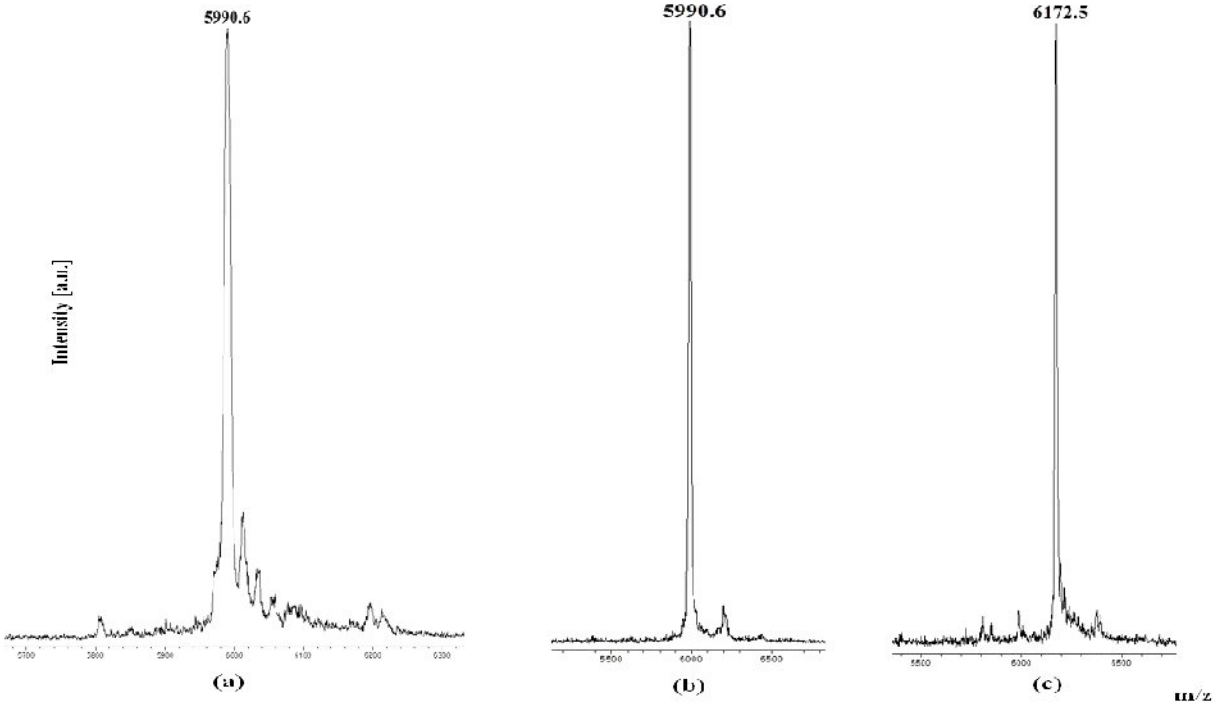


Fig. 3: Identification of dodecanoylated products by MALDI**. (a)**, $[M + 1H]^{+1}$ ion peaks in the crude product at 5990.6 and 6172.5(due to mono- and di-dodecanoyl insulin); **(b)**, at 5990.6 (mono-dodecanoyl insulin, peak1, Fig 2**b**) and **(c),** at 6172.5 (di-dodecanoyl insulin, peak 2, Fig. 2**b**).

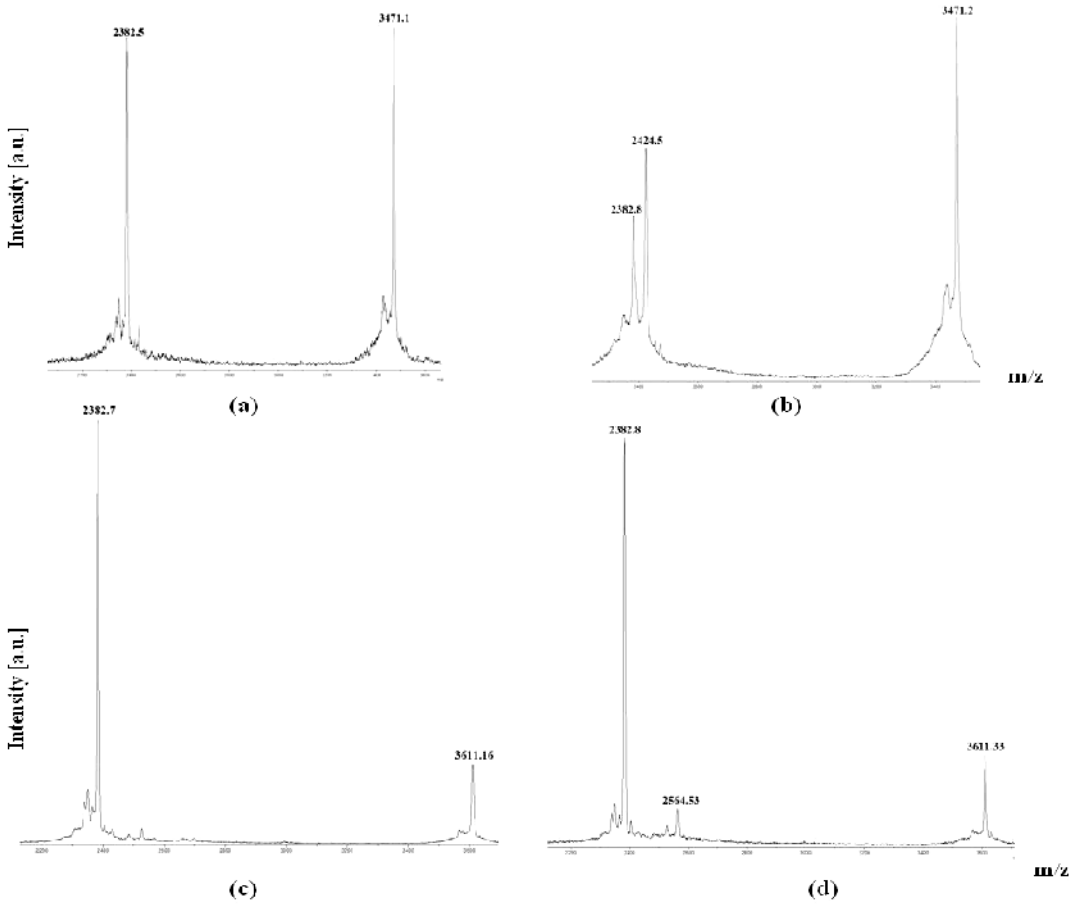


Fig. 4: MALDI mass spectra of thiolytic cleavage products of insulin modified with N-succinimidyl acetate **(a, b)** and dodecanoate (**c, d)** in linear negative mode showing $[M - 1H]^{-1}$ ionic peaks. (**a)**, acetylated-B-chain at 3471.1 and normal A-chain at 2382.5 from mono-acetyl insulin; **(b)**, acetylated-B-chain at 3471.2, normal and acetylated A-chains at 2382.8 and 2424.5 respectively from di-acetyl insulin; (**c),** mono-dodecanoyl B-chain at 3611.16 and normal A-chain at 2382 from mono-dodecanoyl insulin; **(d),** dodecanoyl A and B chains at 2564.5 and 3611.3 respectively from di- dodecanoyl insulin.

It has been observed before, that differences in the pK_a values of α- and ε-amino groups (8.0 and 10.0 respectively) of peptides can influence their reactivity towards acylation reaction. In general here, the modification of insulin by N-succinimidyl acetate at pH 12.0 led to acetylation at the α-amino group of G^{1A} in A-chain and ε-amino group of Lys^{29B} in the B-chain giving rise to an equal mixture of mono- and di-acetylated products. No significant reaction occurred at the amino group of F^{1B} in the B-chain. Since the pK_a values of G^{1A} and $F^{1B}$are likely to be very similar, the lack of modification of the latter may be attributed to steric factors due to a bulky benzyl moiety in the neighborhood of the amino group to be functionalized.

Despite the greatly differing pKa values of α- and ε- amino group of proteins, at pH 12.0 both these will be in a non-protonated and in a reactive form [7]. The high degree of regiospecificity to modification of insulin by N-succinimidyl dodecanoate at pH 12.0, thus, is unlikely to be due to electronic factors and may be attributed to steric factors. Now, the size of the dodecanoyl moiety greatly retarded the modification at the α-amino position because of steric hindrance arising from the peptide chain and the large size of the C_{12} regent, thus predominantly mono-acylated product was observed. This is, despite the presence of a large access of the reagent, but fortunately, rapid hydrolysis of the reagent at high pH, used in the reaction, further contributed to the fact that the rapidly formed initial product,ε- mono-dodecanoyl insulin (4), was protected from further reaction at the G^{1A} amino group.

Using two extreme cases, exemplified by N-succinimidyl acetate and dodecanoate, we have noted

here the contrasting behavior of the acylation reaction depending on the size of the acyl chain. The regiospecificity of the acylation reaction was originally studied using a variety of amino acids and different acyl-transfer reagents [7]. The broad conclusion in the paper was that at high pH values, the acylation reaction is predominantly at the ω-amino group. However, a careful examination of the data [7] shows that the discrimination against the modification of the α-amino group occurred only with bulky acylating reagents. Such a bias was removed when modification involved the transfer of a small acyl moiety. Thus the preference for the modification of ε-amino group of amino acids [7] or of proteins, as shown here, over the α-amino group, at pH values above the pK of both the groups, is governed not by electronic but steric factors. As is the case with the lack of modification at Phe^{B1}which is due to the conventional expectation that the amino group is attached to a secondary C centre, hence more sterically hindered.

In broad conclusion, the results show that from a steric view point, ε-amino group K^{29B} is the least hindered and accessible to C_2 as well as C_{12} reagents, then is the amino group of G^{1A} which is accessible to C_2 but not the C_{12} reagent, finally that of $F^{B1}$is most hindered and accessible to neither.

Acknowledgements

We greatly acknowledge the Higher Education Commission (HEC), Islamabad for funding the purchase, maintenance and repair of mass spectrometers and supporting the various aspects of the present research.

References

1. E. Gershonov, Y. Shechter and M. Fridkin, New concept for long-acting insulin: spontaneous conversion of an inactive modified insulin to the active hormone in circulation: 9-fluorenylmethoxycarbonyl derivative of insulin. *Diabetes*, **48**, 1437 (1999).
2. K. Hinds, J. J. Koh, L. Joss, F. Liu, M. Baudys and S. W. Kim, Synthesis and Characterization of Poly(ethylene glycol)−Insulin Conjugates, *J. Bioconj. Chem*. 11, 195 (2000).
3. M. L. Nucci and A. Abuchowski, The Therapeutic Value of poly(ethylene glycol)-Modified Proteins, *Adv. Drug Deliv. Rev*. **6**, 133 (1991).
4. I. Jonassen, S. Havelund, T. Hoeg-Jensen, D. B. Steensgaard, P. Wahlund and U. Ribel, Design of the Novel Protraction Mechanism of Insulin Degludec, an Ultra-long-Acting Basal Insulin, *Pharm. Res*. **29** 2104 (2012).
5. P. Kurtzhals, S. Havelund, I. Jonassen, B. Kieher, U. D. Larsen, U. Ribel and J. Markussen, Albumin binding of insulins acylated with fatty acids: characterization of the ligand-protein interaction and correlation between binding affinity and timing of the insulin effect *in vivo*. *Biochem. J*. **312**, 725 (1995).
6. S. Havelund, Acylated insulin, U S patent, U S 5750497 A (1998).
7. J. Leclerc and L. Benoiton, On the selectivity of acylation of unprotected diamino acids, *Cand. J. Chem*, **46**, 1047 (1968).
8. Y. S. Lapidot, Rappoport and Y. Wolman, Use of Esters of *N*-hydroxysuccinimide in the Synthesis of *N*-acylamino acids, *J. Lipid Res*, **8**, 142 (1967).
9. Q. A. Gardner, H. Younas and M. Akhtar, Studies on the regioselectivity and kinetics of the action of trypsin on proinsulin and its derivatives using mass spectrometry, *J. Biochem. Biophys.Acta*, **1834**, 182 (2013).
10. F. Ahsan, A. Arif, N. Mahmood, Q. A. Gardner N. Rashid and M. Akhtar, Characterization and bioassay of post-translationally modified ineterferon alpha-2b expressed in *Escherichia coli*, *J. Biotech*. **184**, 11 (2014).