

Investigation of asparagine deamidation in a SOD1-based biosynthetic human insulin precursor by MALDI-TOF mass spectrometry

Anna Bierczyńska-Krzysik^{#1}✉, Małgorzata Łopaciuk¹, Renata Pawlak-Morka² and Dorota Stadnik^{1#}

¹Institute of Biotechnology and Antibiotics, Warszawa, Poland; ²Bioton S.A., Macierzysz, Ozarów Mazowiecki, Poland

A biosynthetic human insulin precursor displayed enhanced susceptibility to deamidation at one particular site. The present study was undertaken to monitor progress of precursor deamidation at successive manufacturing stages. MALDI-TOF/TOF MS in combination with controlled endoproteinase Glu-C and endoproteinase Asp-N proteolysis was used for rapid and unambiguous determination of deamidated residue within the investigated structure. Close inspection of isotopic distribution patterns of peptides resulting from enzymatic digestion enabled determination of distinct precursor forms occurring during the production process. Asn, Asp, isoAsp and succinimide derivatives of the amino acid at position 26 were unambiguously identified. These modifications are related to the leader peptide of a precursor encompassing amino acid sequence corresponding to that of superoxide dismutase [Cu-Zn] (SOD1 1, EC=1.15.1.1). Monitoring of precursor deamidation process at successive manufacturing stages revealed that the protein folding stage was sufficient for a prominent replacement of asparagine by aspartic and isoaspartic acid and the deamidated human insulin precursor constituted the main manufactured product. Conversion proceeded through a succinimide intermediate. Significant deamidation is associated with the presence of SNG motif and confirms results achieved previously on model peptides. Our findings highlight an essential role of the specific amino acid sequence on accelerated rate of protein deamidation. To our knowledge, this is the first time that such a dramatic change in the relative abundance of Asp and isoAsp resulting from protein deamidation process is reported.

Key words: asparagine deamidation; insulin precursor; superoxide dismutase; peptide mass fingerprinting; MALDI-TOF MS

Received: 23 January, 2014; **revised:** 11 May, 2014; **accepted:** 28 May, 2014; **available on-line:** 16 June, 2014

INTRODUCTION

Proteins and peptides are susceptible to spontaneous, nonenzymatic deamidation, as a result of which the amide side-chain of glutamine (Gln) or asparagine (Asn) residues is transformed into a free carboxyl group. Deamidation of Asn residues is more frequent and more rapid than for Gln residues (Wright, 1991; Aswad, 1995; Robinson & Robinson, 2004). Two pathways of the reaction are possible, depending on pH: below pH 3, acid hydrolysis of the amide side chain of Asn residues leads

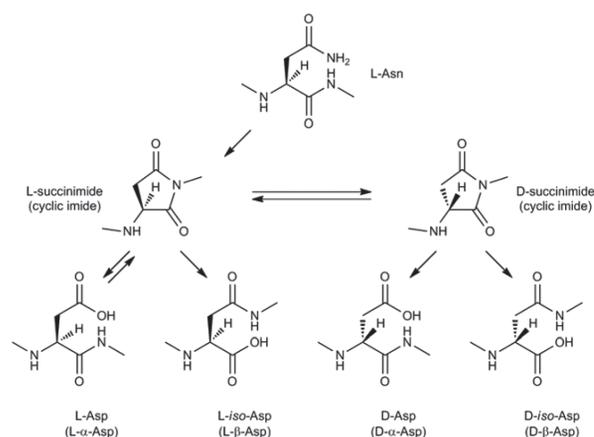


Figure 1. General pathway of deamidation, isomerization and racemization of Asn residue.

to formation of aspartic acid, while for pH above 4, deamidation reaction proceeds through a succinimide intermediate which subsequently hydrolyzes to a mixture of aspartic acid (Asp) and isoaspartic acid (iso-Asp or β-Asp) (Patel & Borchardt, 1990; Brange, 1992; Peters & Trout 2006; Catak *et al.*, 2009) (Fig. 1).

For short peptides without higher order structural elements, predominance of the isoAsp over Asp derivative was observed with approximate ratio of 3:1. However, in case of proteins, favorable changes in conformation might result in a higher relative yield of the Asp derivative (Wearne & Creighton, 1989; Oliyai & Borchardt, 1994; Rivers *et al.*, 2008; Sinha *et al.*, 2009). Since succinimides are prone to racemization, D-aspartyl derivatives can also be produced during hydrolysis (Geiger & Clarke 1987; Li *et al.*, 2003) (Fig. 1).

The conversion of Asn to Asp/isoAsp results in a 1 (0.98) Da mass increase, which can be detected by MS detectors. However, determination whether a protein is deamidated is not straightforward due to overlapping isotopic peaks, unless high resolution MS is used. Usually, identification of deamidation sites in proteins involves protein digestion, separation of the resulting

✉ e-mail: bierczynskaa@iba.waw.pl

[#]These authors contributed equally to this work

Abbreviations: Asn, asparagine; Asp/isoAsp, aspartic acid/isoaspartic acid; SOD, superoxide dismutase; Glu-C, endoproteinase Glu-C (*Staphylococcus aureus* Protease V8); Asp-N, endoproteinase Asp-N; MALDI-TOF, matrix-assisted laser desorption ionization – time of flight analyzer; MS, mass spectrometry

peptides and mass analysis using MS techniques. Distinction of Asp and isoAsp derivatives is even more challenging as isomerization does not imply any mass shift. In some cases the Asp and isoAsp derivatives can exhibit different MS/MS fragmentation patterns, but unfortunately these differences are not specific for all Asp/isoAsp isomers and depend on the protein sequence (O'Connor *et al.*, 2006; Cournoyer *et al.*, 2007). Presence of satellite peaks in mass spectra with a mass shift of -17 Da can be indicative of succinimide intermediate whose existence correlates with formation of Asp/isoAsp isomers and therefore may be useful in identification of the deamidation site. There are also two indirect proofs for the presence of isoAsp residue employed in this study: termination of the Edman degradation reaction at isoAsp and resistance to cleavage by endoproteinase Asp-N (Kameoka *et al.*, 2003). This endoproteinase cleaves peptide bonds on the N-terminal side of Asp but not on isoAsp or Asn residues. Deamidation rate depends on many exogenous (pH, temperature, buffer species, ionic strength) and endogenous (amino acid sequence, secondary and higher order structures) factors. The influence of protein sequence and structure on Asn deamidation is significant. An increased conversion of Asn to Asp/isoAsp derivatives was observed for those proteins in which Asn was followed by small, flexible or hydrophilic residues. In particular, glycine was found to strongly accelerate deamidation rate since this amino acid does not possess a side-chain (Hayes & Setlow, 1997; Aswad, 2000; Chelius *et al.*, 2005; Krokhn *et al.*, 2006). Deamidation of a wild-type SOD-1 protein which contains neighboring residues of asparagine and glycine was first reported very recently (Shi *et al.*, 2013). We have also observed that the determined molecular mass of a recombinant insulin precursor which contains a part of SOD 1 protein sequence is 1 Da higher than that calculated on the basis of the nucleotide sequence. This difference in mass was ascribed to deamidation of one of the Asn residues. Thus, the overall aim of our research was to determine the extent and location of the deamidation site and to investigate if deamidation occurs during downstream processing. In the present study, we employed enzymatic digestion with endoproteinase Asp-N and endoproteinase Glu-C and MALDI-TOF mass spectrometry for these purposes. Our findings undoubtedly indicate that deamidation occurs within a leader peptide of the insulin precursor. In contrast with the paper of Shi *et al.*, 2013, where deamidation reached ~23%, the insulin precursor purified after biosynthesis was nearly completely deamidated.

EXPERIMENTAL SECTION

Reagents and materials. Recombinant DNA derived human insulin precursor was expressed in *E. coli* strain. Acetic acid, ammonium bicarbonate, ammonium hydroxide and ammonia were purchased from POCH S.A. (Gliwice, Poland), trifluoroacetic acid was from Sigma-Aldrich (Munich, Germany), acetonitrile HPLC far UV/gradient grade was from J.T. Baker (Deventer, Holland). Protease from *S. aureus*, strain V8 (endoproteinase Glu-C), originated from MP Biomedicals LLC (cat. no. 151972) and Endoproteinase Asp-N from Sigma Aldrich (cat. no. P3303). Both enzymes were used for digestion of human insulin precursor. Alpha-cyano-4-hydroxycinnamic acid (CHCA) from Fluka (cat. no. 70990), was exploited as a MALDI-MS matrix. Ultrapure water from

NANOpure Diamond Ultrapure Water Systems (Barnstead International, Boston, USA) was used in all experiments.

Sample preparation. Purified precursor. Insulin precursor digestion using endoproteinase Glu-C at pH 8.

The precursor was dissolved in 0.1 M NH_4HCO_3 to concentration of 1 mg/ml and pH 8.0 ± 0.2 . Digestion was carried out using endoproteinase Glu-C with enzyme to protein ratio 1:10 (w/w), temperature 37°C and rotation speed 350 rpm for 24 h.

Insulin precursor digestion using endoproteinase Glu-C at pH 4. All samples were diluted with 0.1 M CH_3COOH (pH 4.0) to concentration of 1 mg/ml and pH 4.0 ± 0.2 . Digestion was carried out using endoproteinase Glu-C with enzyme to protein ratio 1:10 (w/w), temperature 37°C and rotation speed 350 rpm for 24 h.

Insulin precursor digestion using endoproteinase Asp-N at pH ~8. All samples were diluted with 0.1M NH_4HCO_3 to concentration of 1 mg/ml and pH 8.0 ± 0.2 . Digestion was carried out using endoproteinase Asp-N with enzyme to protein ratio 1:50 (w/w), temperature 37°C and rotation speed 350 rpm for 24 h. MS analysis followed after 4 h and 22 h of digestion.

Precursor samples collected at different stages of downstream processing: dissolution of the inclusion bodies, beginning and end of folding. The protein concentration in samples was determined with the use of Bradford assay. During downstream processing two samples were collected from each stage. First one was adjusted to pH 3, the second one remained unchanged. Samples from each production stage were digested independently with the use of endoproteinase Glu-C and endoproteinase Asp-N. Both enzymes are active at alkaline pH (pH ~ 8). Endoproteinase Glu-C is also active in acidic environment (pH ~ 4). The digestions with endoproteinase Glu-C (pH 4 and 8) and endoproteinase Asp-N (pH 8) were performed in an analogous way for the purified precursor. MS analysis followed after 4 h and 22 h of digestion.

Mass spectrometry. MALDI-TOF MS experiments were performed in a reflector mode with the use of 4800 Plus (AB SCIEX) instrument. CHCA matrix was exploited. External calibration was achieved with a peptide mixture provided by AB SCIEX (4700 Cal Mix). Samples were purified with Zip-Tip C18 (Millipore) and air-dried. Each sample was applied 10 times on a MALDI-TOF target plate and analyzed. Experiment involving analysis of samples at successive manufacturing stages was repeated with three batches of bacteria.

RESULTS AND DISCUSSION

Identification of the asparagine deamidation site in biosynthetic human insulin precursor by MALDI-TOF mass spectrometry

Human insulin precursor, produced in *E. coli*, consists of the leader peptide and the amino acid sequence of the proinsulin in the following configuration: leader peptide-B-peptide C-A, where A and B correspond to the human insulin chains. The amino acid sequence of the leader peptide is similar to that of Superoxide dismutase [Cu-Zn] (EC 1.15.1.1, Accession No.; P00441) (2-62), with the exception that all the cysteines were replaced with serines. During insulin production the leader peptide and peptide C are further removed by enzymatic digestion to obtain native human insulin hormone.

10 20 30 40 50 60
 ATKAVSVLKG DGPVQGIINF EQKESNGPVK VWGSIKGLTE GLHGHHVHEF GDNTAGSTSA

Figure 2. Predicted sequence of the first 60 amino acids of the leader peptide.

The numbers above the sequence presented reflect the order of appearance of particular amino acids. The whole sequence of the precursor is not shown because of its commercial interest.

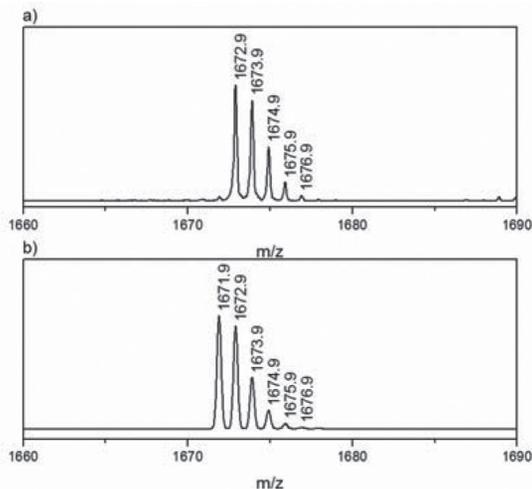


Figure 3. MALDI-TOF mass spectrum of peptide 25–40 resulting from endoproteinase Glu-C digestion of purified human insulin precursor at pH 8:

(a) observed, (b) predicted isotopic distribution pattern for the peptide SNGPVKVVWGSIKGLTE with the use of mMass, an open-source MS tool.

The amino acid sequence of the leader peptide predicted from the cDNA should be as presented in Fig. 2.

Surprisingly, Edman sequencing performed on the purified human insulin precursor stopped at position 26 (not shown). In order to further explore the suspected residue, purified precursor was subjected to 4 h of pro-

teolysis with endoproteinase Glu-C at pH 8. The peptide mixture was then analyzed by MALDI-TOF with particular attention paid to peptides containing amino acid 26. The determined m/z value of the peptide 25–40 resulting from the digest revealed that the peptide mass is nearly 1 Da higher than expected (Fig. 3a). Atypical isotopic distribution of the peptide 25–40, SNGPVKVVWGSIKGLTE, was compared to a mass spectrum predicted by the mMass (an open-source MS tool), for the peptide 25–40 with consideration of asparagine presence at position 26 (Fig. 3b). The difference in relative intensity of the monoisotopic ions peak at m/z 1671.9 and 1672.9 is significant. The above results and termination of Edman sequencing at position 26 suggest that position 26 may indeed contain isoaspartic acid and not asparagine as in the predicted amino acid sequence. To exclude possibility of a point mutation, nucleotide sequence of the human insulin precursor gene was verified by DNA sequencing. The results confirmed the presence of nucleotides coding the Asn residue. In consequence, the observed mass shift led to the supposition that the precursor might be fully deamidated at this specific position.

In order to confirm the fragment sequence, the peptide at m/z 1672.9 was subjected to MS/MS analysis. The spectrum obtained is given in Fig. 4.

The MS/MS spectra were submitted to the Mascot search engine (Matrix Science Ltd., London) to identify the peptide sequence. Results retrieved from the Mascot database search clearly indicated the presence of aspartic/isoaspartic acid and not asparagine at position 26. Mascot results do not differentiate between Asp and isoAsp.

Due to the fact that the digestion was carried out at basic pH, which could induce protein deamidation during sample handling (Krokhin *et al.*, 2006), proteolysis with endoproteinase Glu-C was also performed at pH 4 and 2.8. MALDI-TOF measurements followed after 4 and 22 h after enzyme addition. All mass spectra recorded strongly resembled the one presented in Fig. 3a, suggesting that the obtained purified human insulin pre-

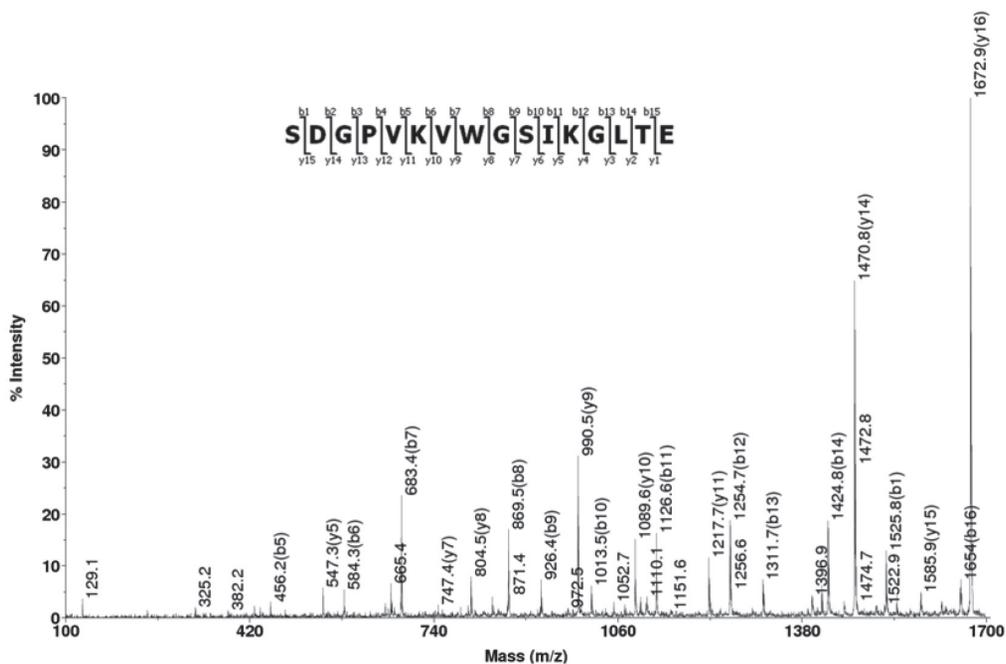


Figure 4. Exemplary MS/MS spectrum of the SDGPVKVWGSIKGLTE peptide, acquired with MALDI-TOF/TOF instrument, resulting from fragmentation of the peak at m/z 1672.9.

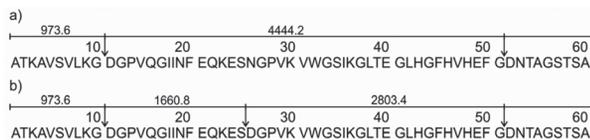


Figure 5. Proteolysis of the purified human insulin precursor with endoproteinase Asp-N:

(a) position 26 occupied by asparagine, (b) position 26 occupied by aspartic acid; the arrows indicate peptide bonds susceptible to digestion. Theoretical $[M+H]^+$ values of resulting peptides are depicted above the peptide sequence.

cursor was already deamidated at position 26 prior to proteolytic digestion and mass measurements.

Since downstream processing of the precursor proceeds at strong alkaline pH, the resulting deamidated product can contain either isoAsp or Asp residue. To determine which isomer was formed here, proteolytic digestion with endoproteinase Asp-N and MALDI-TOF mass analysis were applied.

The precursor possessing asparagine at position 26, cleaved with endoproteinase Asp-N, should undergo disruption in two places, indicated with arrows in Fig. 5a.

The same peptide bonds are to be hydrolyzed in the presence of isoaspartic acid at position 26. In this case, a 1 Da shift in m/z value of resulting peptides could serve as an indicator of amino acid composition at this position. If aspartic acid was present at position 26, three peptide bonds would be interrupted, as illustrated with arrows in Fig. 5b.

MALDI-TOF mass spectrum of the precursor digested with endoproteinase Asp-N (Fig. 6) displayed characteristic peaks at m/z 973.6, 4445.2 corresponding to the peptides 1–10 and 11–51 respectively. The peptides result from the precursor disruption at two sites, however the ion peak at m/z 4445.2 exhibits a mass increase of 1 Da, strongly indicating the existence of deamidation in the proteolytic peptide 11–51. The isotopic distribution pattern of this ion peak is in agreement with the one visualized by the MS-Isotope tool for the peptide 11–51 containing isoAsp at position 26 (data not shown). Addi-

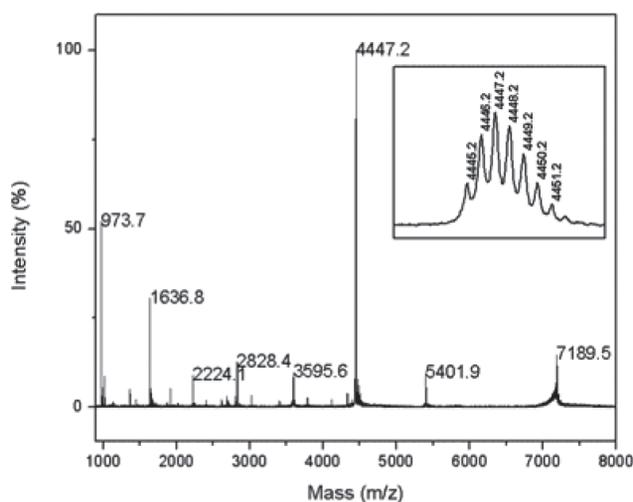


Figure 6. Mass spectrum of purified human insulin precursor subjected to 22 h of proteolysis with endoproteinase Asp-N at pH 8.

Characteristic peaks at m/z 973.6, 4445.2 correspond to 1–10 and 11–51 peptides respectively. The inset shows close inspection of the signal at m/z 4445.2. Remaining signals do not correspond to the peptides containing residue 26 and therefore were excluded from further investigation.

tionally, the absence of signals at m/z 1660.8 and 2803.4 (indicative of Asp presence at position 26) is likely to be caused by the modification of aspartate residue at position 26 and the inability of endoproteinase Asp-N to cleave N-terminal site of the isomerized aspartyl residue.

Resulting mass spectra and nearly 1 Da mass shift in m/z value of peptides containing amino acid 26 allowed concluding that the analyzed protein contained mainly isoaspartic acid at position 26. The lack, or very low intensity of signals at m/z 1660.82 (theoretical value, 11–25; DGPVQGIINFEQKES) and 2803.4 (theoretical value, 26–51; DGPVKVWGSIKGLTEGLHGF HVHEFG) predicted for peptides containing aspartic acid at position 26, indicates that the majority of this amino acid is present in its β form. This finding is also confirmed by termination of Edman sequencing at position 26 (not shown).

Reassuring, based on Edman sequencing results and careful investigation of isotopic distribution pattern of peptides resulting from enzymatic cleavage of purified human insulin precursor, it was found that there is a considerable domination of precursor derivative containing isoaspartic acid over asparagine and aspartic acid at position 26. The conversion of asparagine to isoaspartic acid and aspartic acid, through the cyclic imide intermediate, already occurs in the production or purification process and is not a consequence of sample handling directly prior and during Edman sequencing, proteolysis or MALDI-MS.

Monitoring of asparagine deamidation in human insulin precursor during manufacturing process

In the light of the results described above and the divergence between the nucleotide sequence and the experimentally established amino acid composition of the purified human insulin precursor, more detailed analysis of the isotopic distribution patterns of peptides containing amino acid no. 26 on successive stages of the insulin production (Fig. 7) was undertaken. The study was performed with three batches of bacteria. Precursor samples were collected at three production stages: after dissolution of inclusion bodies (2), at the beginning (3) and at the end of folding (4) (Fig. 7). These production stages proceed at strong alkaline pH, which is about one unit higher for dissolution than for folding.

Each sample was divided into two portions, one of which was kept at unchanged alkaline pH and the other was immediately acidified in order to evaluate the eventual role of a high pH value on protein deamidation during storage. Collected samples were subjected to enzymatic digestion with endoproteinase Glu-C at pH 4 and 8 and endoproteinase Asp-N. The resulting fragments were analyzed with MALDI-TOF MS. To simplify the description of numerous mass spectra analyzed, theoretical values for monoisotopic, singly charged signals were employed: 1672.9 to describe peptide SDGPVKVWGSIKGLTE containing aspartic and/or isoaspartic acid at position 26 of the precursor, 1671.9 for peptide SNGPVKLVWGSIKGLTE possessing asparagine and 1654.9 for the intermediate derivative — succinimide.

The analyzed samples constituted a specific mixture of precursor molecules, containing asparagines, aspartic acid and isoaspartic acid

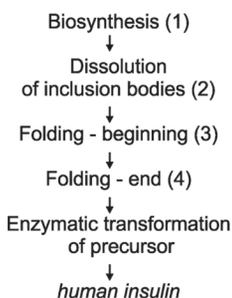


Figure 7. General scheme of recombinant human insulin manufacturing process.

The numbers assigned to successive stages are further quoted in the paper.

at position 26, in ratios dependent on the production stage. Therefore, isotopic distribution patterns of the resulting peptides were atypical. In particular, the diminished intensity of the monoisotopic ion 1671.9, in comparison to the simulated profile, was clearly visible (Fig. 3), suggesting that the analyzed peptide is in fact a mixture of three fragments — Asn, Asp and isoAsp at position 26. Even though the applied MALDI-MS analysis can not be considered as quantitative,

the high reproducibility of results obtained for the analyzed samples enables determination of changes in the measured signal intensities and thus conclusions can be made on the content of individual derivatives in analyzed samples.

In order to refer to the degree of asparagine deamidation at successive production stages, a simulation of isotopic distribution patterns for the peptide SNG-PVKVWGSIKGLTE, with variable content of asparagine/aspartic acid at position 26, was performed with the use of mMass, open-source MS tool. The resulting spectra are presented in Fig. 8. As described, the relative intensity of the first and second isotopic peaks changes according to the progress of peptide deamidation. Maximum monoisotopic peak intensity for peptides containing exclusively asparagine at position 26 occurs at m/z 1671.9, whereas if only aspartic/isoaspartic acid is present, the peak shifts to m/z 1672.9. In the case of an equimolar peptide mixture of SNG-PVKVWGSIKGLTE and SDGPVKVWGSIKGLTE, the intensity of the 1671.9 m/z signal should reach half of intensity of the 1672.9 m/z signal. To summarize, the extent of deamidation can be estimated on the basis of relative intensity of the 1671.9 and 1672.9 peaks in the isotopic distribution pattern.

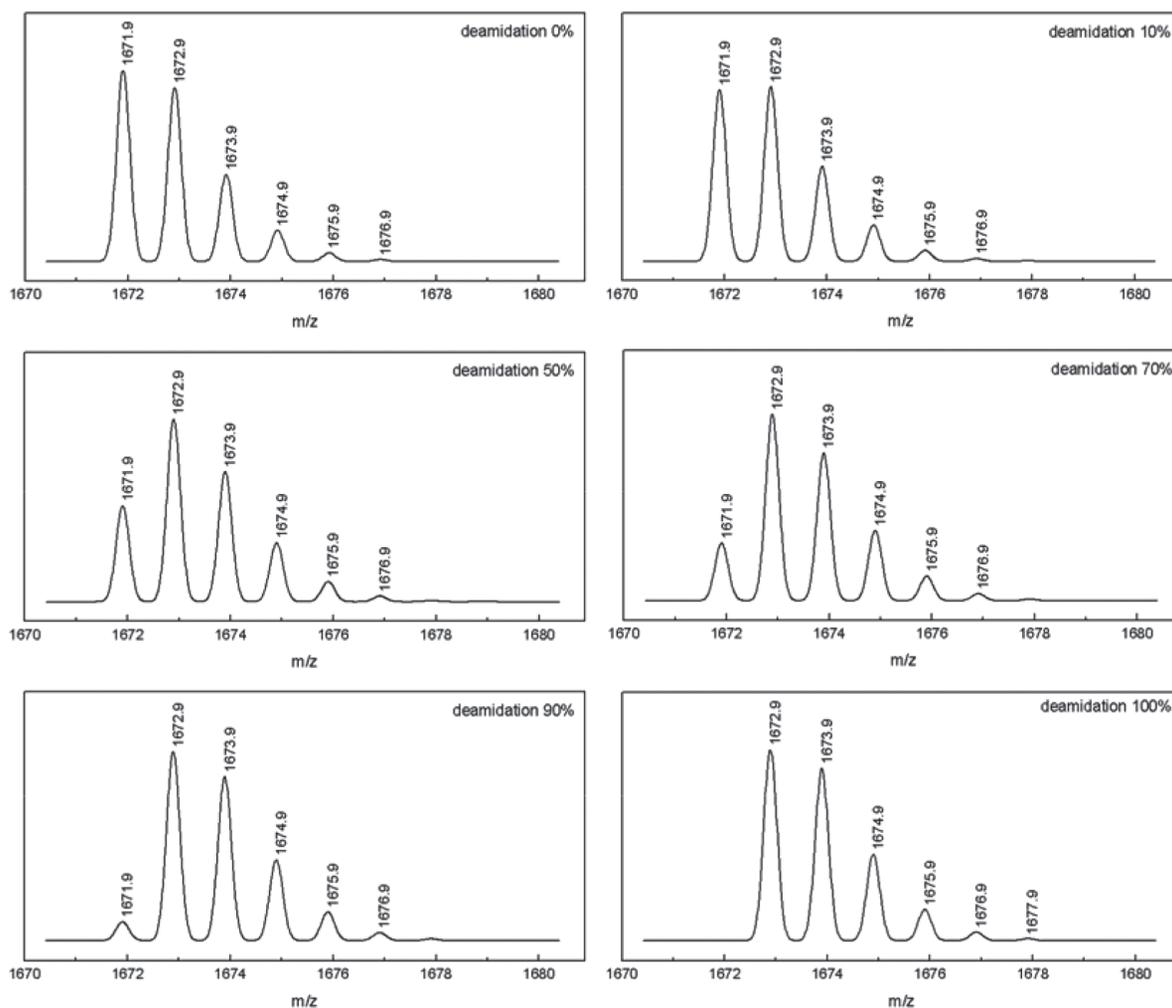


Figure 8. Isotopic distribution pattern of SXGPVKVWGSIKGLTE peptide — a computer simulation using mMass, open-source MS tool. The X stands for Asn, Asp or isoAsp depending on the deamidation extent.

Human insulin precursor digestion with endoproteinase Glu-C at alkaline and acidic pH; samples acidified immediately after collection

Directly after collection, samples from stages 2, 3 and 4 were acidified to pH 3 and stored frozen prior to further investigation. Successively, samples were digested with endoproteinase Glu-C at pH 8 and 4 and the resulting fragments were analyzed with MALDI-TOF MS. Close examination of the spectrum of the precursor from stage 2, digested at acidic and alkaline pH, revealed a slight predominance of the signal at m/z 1671.9 over that of 1672.9 for the acidic digest (Fig. 9 b, stage 2). An opposite difference in the intensity of those two peaks was observed for the alkaline digest. (Fig. 9 a, stage 2). Analysis of further mass spectra (Fig. 9 a and b, stage 4) revealed a clear decrease in the intensity of the signal at m/z 1671.9, characteristic for the peptides containing asparagine at position 26. The domination of the signal at m/z 1672.9 over that of 1671.9 is obviously most significant for precursor samples collected after folding. Furthermore, the intensity difference of these two signals is higher for the sample cleaved at pH 8.

The presence of a peak at m/z 1654.9 in the analyzed samples, distinctive for succinimide, confirms the mechanism of deamidation depicted in Fig. 1. The intensity of the signal at m/z 1654.9 remains at a comparable level for samples taken from stage 2 and 3 and digested at pH 8. A noticeable drop in the signal intensity is observed for samples collected at the end of folding (stage 4), suggesting that most of asparagine residues have been already converted to aspartic/isoaspartic acid. In contrast, for the acidic digest, the peak of succinimide becomes only visible in the sample from stage 4. This finding suggests that the deamidation and isomerization are restrained here compared to the samples maintained for a longer time at alkaline pH. Therefore, to ascertain the level of Asn/Asp/isoAsp in the untreated, original samples, storage and digestion should be performed at acidic pH. Maintenance of low pH prevents deamidation via succinimide intermediates.

Based on the obtained results, it is concluded that at stage 2 (the dissolution of inclusion bodies), and 3 (the beginning of folding), it is mostly asparagine which occupies position 26 of the insulin precursor. In the subsequent manufacturing stages, due to prolonged exposure to high pH, asparagine is gradually converted to aspartic and isoaspartic acid. The key manufacturing stage, leading to significant asparagine deamidation, is the folding process. The further progress of asparagine deamidation probably occurs at succeeding purification stages (stage 5 in Fig. 7).

Human insulin precursor digestion with endoproteinase Glu-C at alkaline and acidic pH; original, alkaline samples

Similar analyses were carried out for precursor samples collected from manufacturing stages 2, 3 and 4 (shown in Fig. 7) and stored at alkaline environment prior to digestion with endoproteinase Glu-C at pH 8 and 4. In both cases, extensive deamidation of peptide SNGPVKVVWGSIKGLTE was observed. Analysis of mass spectra registered for the sample from stage 2, after 4 h of cleavage, allowed the observation of domination of a signal at m/z 1672.9, corresponding to the SDGPVKVVWGSIKGLTE peptide containing aspartic/isoaspartic acid at position 26 (Fig. 10a and b). Moreover, it is possible to distinguish a signal at a 1 Da lower m/z value, which constitutes a proof of parallel presence

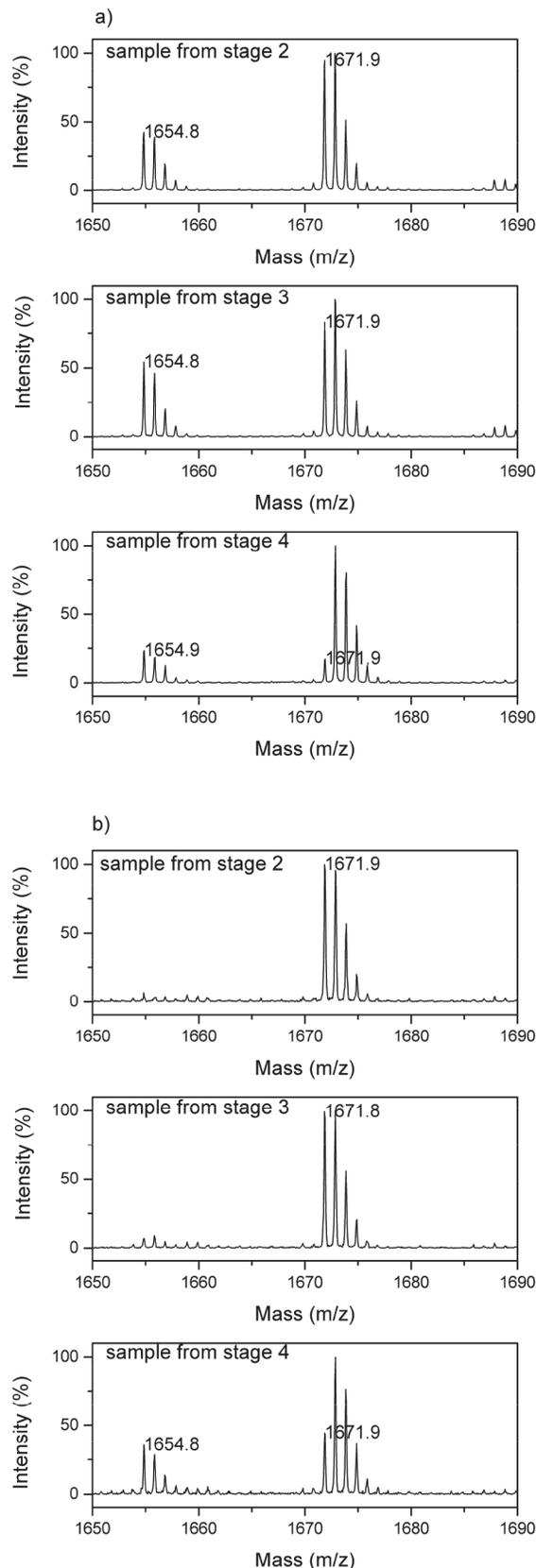


Figure 9. Mass spectra fragments resulting from human insulin precursor cleavage with endoproteinase Glu-C at alkaline (a) and acidic (b) pH (4 h of proteolysis), performed on acidified samples from stages 2, 3 and 4. Monoisotopic peaks were labeled.

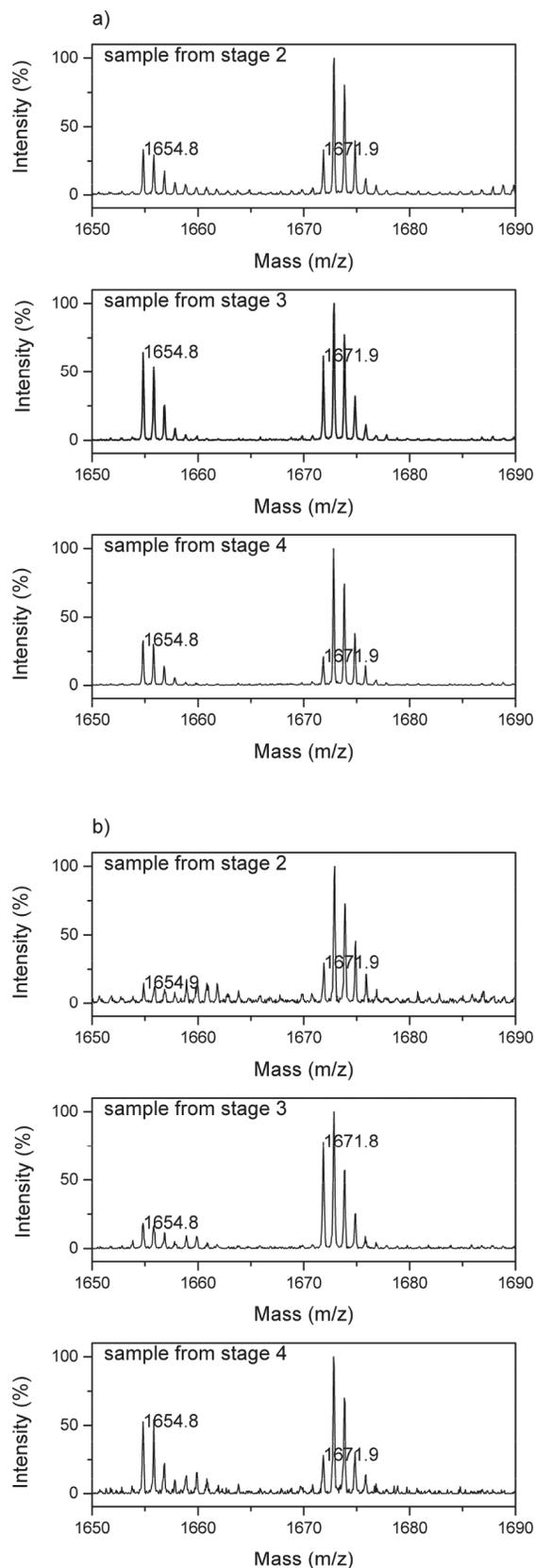


Figure 10. Mass spectra fragments resulting from human insulin precursor cleavage with endoproteinase Glu-C at alkaline (a) and acidic (b) pH (4 h of proteolysis), performed on samples from stages 2, 3 and 4. Monoisotopic peaks were labeled.

of asparagine in the digest. Slight lowering of pH at the beginning of folding (stage 3) results in increased intensity of the 1671.9 *vs.* 1672.9 signal, compared to peak intensity relations from stage 2. However, it is still the 1672.9 signal that dominates. After folding (stage 4) the predominance of the 1672.9 over 1671.9 signal, in other words aspartic/isoaspartic acid over asparagine at position 26, is most evident. However, great care must be taken when drawing conclusions from these data. Current storage conditions (alkaline pH) of collected precursor samples and digestion protocols inevitably cause artificial deamidation and do not reflect composition of the original samples. Thus, the rate of precursor deamidation at successive manufacturing stages cannot be estimated on the basis of data for samples stored and digested at high pH. Nevertheless, interesting observations can be made from comparison of the mass spectra shown in Fig. 9 and 10: the isotope distribution patterns of the acidic digests (Fig. 9b and Fig. 10b) are different. The only difference between these two digests was storage condition of the original samples prior to analysis. The samples, for which mass spectra are shown in Fig. 9b, were acidified after collection, whereas the pH of samples for which mass spectra are summarized in Fig. 10b was kept high. Both samples were frozen at -20°C prior to proteolysis. This suggests that deamidation noticeably progressed at alkaline environment even at low temperature.

Human insulin precursor digestion with endoproteinase Asp-N

Combination of mass spectrometry and enzymatic cleavage with endoproteinase Asp-N was applied in an

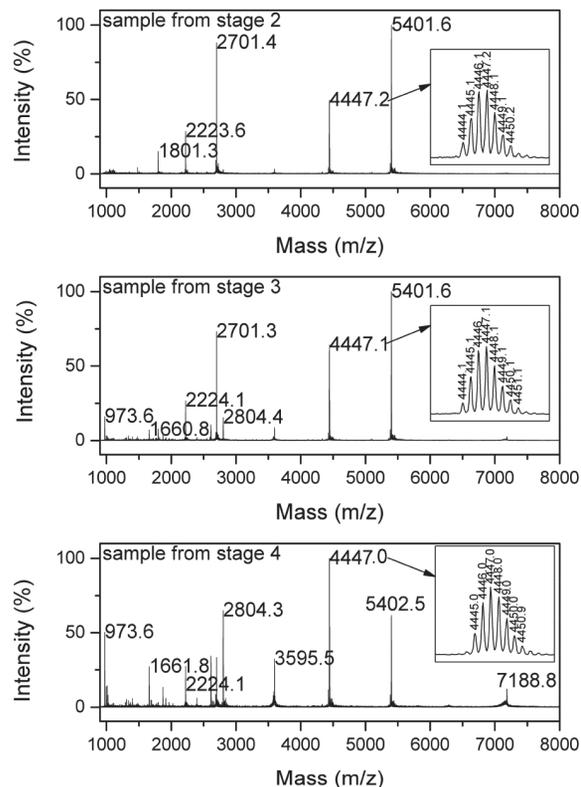


Figure 11. Mass spectra fragments resulting from human insulin precursor cleavage with endoproteinase Asp-N at alkaline pH (4h of proteolysis), performed on acidified samples from stages 2, 3 and 4.

attempt to distinguish between Asp and isoAsp derivatives of precursor at different manufacturing stages. For these purposes, the precursor samples from manufacturing stages 2, 3 and 4, acidified after collection, were digested with endoproteinase Asp-N at pH 8 and analyzed by MALDI-TOF MS. To simplify description of the mass spectra obtained, theoretical values for monoisotopic, singly charged signals were employed: 4444.2 to describe peptide DGPVQGIINFEQKESNGPVKVV-GSIKGLTEGLHGFHVHEFG containing asparagine at position 26 of the precursor and 4445.2 for DGPVQGIINFEQKESDGPVKVWGSIKGLTEGLHGFHVHEFG possessing isoaspartic acid.

Careful examination of the registered mass spectra presented in Fig. 11, revealed the gradual decrease of the 4444.2 m/z signal in reference to the 4445.2 signal in the subsequent manufacturing stages.

The latter signal represents a peptide containing isoaspartic acid at position 26 of the human insulin precursor. Due to the fact that the analyzed samples do not contain a purified precursor but constitute a specific mixture of precursor molecules containing at position 26 asparagine, aspartic acid and isoaspartic acid, in ratios dependent on the production stage, the resulting mass spectra are very heterogeneous and complex for interpretation. The presence of aspartic acid at position 26 can be confirmed by the presence of 1660.8 and 2803.4 signals in the endoproteinase Asp-N digest spectrum. The values correspond to peptides DGPVQGIINFEQKES (11–25) and DGPVKVWGSIKGLTEGLHGF HVHEFG (26–51), respectively.

CONCLUSIONS

We have demonstrated that high pH applied during manufacturing of recombinant human insulin precursor leads to unintentional protein deamidation at one specific position within the leader peptide. In the case of the presented protein sequence fragment, modification is developed with a very high rate and leads to a nearly complete conversion of asparagine to aspartic and isoaspartic acid. As a result of research performed both on the purified human insulin precursor as well as several manufacturing stages, we have shown that purified insulin precursor does not contain asparagine at position 26, as coded by the nucleotide sequence, but its deamidated counterparts — aspartic and isoaspartic acid. Insulin precursor at earlier manufacturing stages is, in fact, a mixture of derivatives containing asparagine, aspartic and isoaspartic acid at position 26 as well as the intermediate form - succinimide.

The insulin precursor contains six asparagine residues but only the amino acid no. 26 constitutes a “hot spot” for fast, spontaneous deamidation. Presence of glycine and serine as neighboring amino acids of asparagines no. 26 (SNG) in the precursor sequence may enhance its conversion. The leader peptide also has an NT and a DG sequences (prone to deamidation) but no deamidated or isomerized species were found at these sites. Moreover, commercially available SOD 1 protein submitted to long exposure time at high pH, according to the insulin precursor manufacturing process (Fig. 6, stages 2, 3 and 4), revealed the same tendency for deamidation like the part of its sequence present in recombinant human insulin precursor (not shown).

The studies conducted on samples collected from consecutive production stages undeniably demonstrate the progressive increase of aspartic and isoaspartic acid con-

tent at the expense of asparagine. Research performed on isolated and purified insulin precursor, in compliance with Edman degradation results, revealed the dominating presence of isoaspartic acid at position 26. In the work described by Shi who first reported deamidation of a SOD 1 protein (Shi *et al.*, 2013), isoAsp remained undetectable. Asparagine is visible at the inclusion bodies dissolution stage, whereas in the final stage, the amino acid can hardly be detected. Deamidation is promoted by alkaline pH. The folding stage is sufficient for a prominent replacement of asparagine by aspartic and isoaspartic acid at this one specific position.

It is important to note that the SNG sequence is present in naturally occurring SOD1. This enzyme has attracted significant attention because of its antioxidant and pharmacological properties as an anti-inflammatory agent or a remedy for ischemic cardiac diseases (Yasui & Baba 2006; Schaller *et al.*, 2012; Carillon *et al.*, 2013). It can be expressed at high yield in recombinant bacteria (Kang *et al.*, 1997) and yeast (Li & Yu, 2007), therefore the SOD1 gene is often fused to other proteins' genes to improve expression efficiency. For that reason it is important to detect possible post-translational modifications, especially for recombinant protein drug candidates, for which any discrepancies between nucleotide and amino acid sequences should be explained and described in a dossier for submission to regulatory authorities. To our knowledge, in spite of numerous publications referring to SOD1, including PTM modifications (Zhang *et al.*, 2004; Furukawa & O'Halloran, 2006; Wilcox *et al.*, 2009), such extensive deamidation of asparagine at position 26 and formation of its isoAsp derivative has not been described yet. Our observation is particularly relevant for SOD1-based medicines as deamidated forms may show different activity. One should be aware that high pH during any stage of manufacturing may enhance deamidation of SOD1 at this position.

Acknowledgements

We are indebted to Michał Odrowąż-Sypniewski, our lab colleague, for his technical assistance.

REFERENCES

- Aswad DW (1995) Deamidation and isoaspartate formation in proteins. In Aswad DW, ed. CRC Press, Boca Raton.
- Aswad DW (2000) Isoaspartate in peptides and proteins: formation, significance, and analysis. *J Pharm Biomed Anal* **21**: 1129–1136.
- Brange J (1992) Chemical stability of insulin. 4. Mechanisms and kinetics of chemical transformations in pharmaceutical formulation. *Acta Pharm Nord* **4**: 209–222.
- Carillon J, Rouanet JM, Cristol JP, Brion R (2013) Superoxide dismutase administration, a potential therapy against oxidative stress related diseases: several routes of supplementation and proposal of an original mechanism of action. *Pharm Res* **30**: 2718–2728.
- Catak S, Monard G, Aviyente V, Ruiz-Lopez MF (2009) Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms. *J Phys Chem A* **113**: 1111–1120.
- Chelius D, Rehder DS, Bondarenko PV (2005) Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. *Anal Chem* **77**: 6004–6011.
- Cournoyer JJ, Lin C, Bowman MJ, O'Connor PB (2007) Quantitating the relative abundance of isoaspartyl residues in deamidated proteins by electron capture dissociation. *J Am Soc Mass Spectrom* **18**: 48–56.
- Furukawa Y, O'Halloran TV (2006) Posttranslational modifications in Cu,Zn-superoxide dismutase and mutations associated with amyotrophic lateral sclerosis. *Antioxid Redox Signal* **8**: 847–867.
- Geiger T, Clarke S (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *J Biol Chem* **262**: 785–794.
- Hayes CS, Setlow P (1997) Analysis of deamidation of small, acid-soluble spore proteins from *Bacillus subtilis* *in vitro* and *in vivo*. *J Bacteriol* **179**: 6020–6027.

- Kameoka D, Ueda T, Imoto TA (2003) Method for the detection of asparagine deamidation and aspartate isomerization of proteins by MALDI/TOF-mass spectrometry using endoproteinase Asp-N. *J Biochem* **134**: 129–135.
- Kang JH, Choi BJ, Kim SM (1997) Expression and characterization of recombinant human Cu, Zn-superoxide dismutase in *Escherichia coli*. *J Biochem Mol Biol* **30**: 60–65.
- Krokhin OV, Antonovici M, Ens W, Wilkins JA, Standing KG (2006) Deamidation of -Asn-Gly- sequences during sample preparation for proteomics: consequences for MALDI and HPLC-MALDI analysis. *Anal Chem* **78**: 6645–6650.
- Li B, Borchardt RT, Topp EM, VanderVelde D, Schowen RL (2003) Racemization of an asparagine residue during peptide deamidation. *J Am Chem Soc* **125**: 11486–11487.
- Li JR, Yu P (2007) Expression of Cu, Zn-superoxide dismutase gene from *Saccharomyces cerevisiae* in *Pichia pastoris* and its resistance to oxidative stress. *Appl Biochem Biotechnol* **136**: 127–39.
- O'Connor PB, Cournoyer JJ, Pitteri SJ, Chrisman PA, McLuckey SA (2006) Differentiation of aspartic and isoaspartic acids using electron transfer dissociation. *J Am Soc Mass Spectrom* **17**: 15–19.
- Oliyai C, Borchardt RT (1994) Chemical pathways of peptide degradation. VI. Effect of the primary sequence on the pathways of degradation of aspartyl residues in model hexapeptides. *Pharm Res* **11**: 751–758.
- Patel K, Borchardt RT (1990) Chemical pathways of peptide degradation. II. Kinetics of deamidation of an asparaginyl residue in a model hexapeptide. *Pharm Res* **7**: 703–711.
- Peters B, Trout BL (2006) Asparagine deamidation: pH-dependent mechanism from density functional theory. *Biochemistry* **45**: 5384–5392.
- Rivers J, McDonald L, Edwards IJ, Beynon RJ (2008) Asparagine deamidation and the role of higher order protein structure. *J Proteome Res* **7**: 921–927.
- Robinson NE, Robinson AB (2004) Deamidation of asparaginyl and glutaminyl residues in peptides and proteins. In *Molecular Clocks*. Alt-house Press, Cave Junction OR.
- Schaller G, et al. (2012) Human pharmacokinetics of intravenous recombinant human Cu/Zn superoxide dismutase. *Int J Clin Pharmacol Ther* **50**: 413–417.
- Shi, Y, et al. (2013) Deamidation of asparagine to aspartate destabilizes Cu, Zn superoxide dismutase, accelerates fibrillization, and mirrors ALS-linked mutations. *J Am Chem Soc* **135**: 15897–15908.
- Sinha S, Zhang L, Duan S, Williams TD, Vlasak J, Ionescu R, Topp EM (2009) Effect of protein structure on deamidation rate in the Fc fragment of an IgG1 monoclonal antibody. *Protein Sci* **18**: 1573–1584.
- Wearne SJ, Creighton TE (1989) Effect of protein conformation on rate of deamidation: ribonuclease A. *Proteins* **5**: 8–12.
- Wilcox KC et al. (2009) Modifications of superoxide dismutase (SOD1) in human erythrocytes: a possible role in amyotrophic lateral sclerosis. *J Biol Chem* **284**: 13940–13947.
- Wright HT (1991) Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. *Crit Rev Biochem Mol Biol* **26**: 1–52.
- Yasui K, Baba A (2006) Therapeutic potential of superoxide dismutase (SOD) for resolution of inflammation. *Inflamm Res* **55**: 359–63.
- Zhang H, Joseph J, Crow J, Kalyanaraman B (2004) Mass spectral evidence for carbonate-anion radical induced posttranslational modification of tryptophan to kynurenine in human Cu,Zn superoxide dismutase. *Free Radic Biol Med* **37**: 2018–2026.