Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Pharmaceutical and Biomedical Analysis 40 (2006) 185–189
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Short communication

Method for determining the average degree of substitution of *o*-vanillin derivatized porcine somatotropin

Thomas R. Sharp *, Ronald Morris, George J. Horan, Linda H. Pezzullo, Justin G. Stroh

Analytical Research & Development Department, mailstop 8286-03, Pfizer Global Research & Development, Eastern Point Road, Groton, CT 06340, USA

Received 10 May 2005; received in revised form 17 June 2005; accepted 22 June 2005
Available online 3 August 2005

Abstract

Electrospray mass spectral observation directly on a sample of a derivatized protein, such as porcine somatotropin (pST), affords a method for evaluating the degree of substitution of this protein. Derivatization of the lysine residues and the terminal amino residue here by formation of a Schiff base with a small aromatic aldehyde (in this case, *o*-vanillin) affords stabilization of the protein so that it may be used in a controlled release veterinary pharmaceutical formulation. This method permits direct observation of substitutions, optimization of manufacturing procedures for producing a commercial product, and permits quality evaluation of material.

Keywords: *o*-Vanillin; Porcine somatotropin; Protein derivatization; Average degree of substitution measurement

1. Introduction

Stabilization of a protein to long-term storage under ordinary conditions – i.e., as a dry powder, at room temperature – is a desirable, if not essential, goal for producing a viable commercial pharmaceutical product [1]. Requiring sub-ambient temperature storage between manufacture and delivery to the patient is impractical, especially for veterinary protein pharmaceuticals. One method being explored is the derivatization of the lysine ε-amino residues in a protein through formation of a Schiff base with a small aromatic aldehyde (Scheme 1). Derivatization of porcine somatotropin (pST), a 20 kDa protein, with *o*-vanillin converts a protein requiring storage at −70 °C to one which can be stored as a powder at room temperature. To evaluate the process for preparing derivatized protein, however, a method is needed to measure the degree of substitution. Electrospray mass spectrometry presents a direct method for measuring the number-average degree of substitution of *o*-vanillin-derivatized porcine somatotropin (OV-pST). The method should be generalizable for determination of average degree of substitution of any variably derivatized protein. The example porcine somatotropin was under development in a sustained release formulation to promote lean meat development in pigs destined for consumption.

2. Experimental

Methods for manufacture of derivaterized protein have been fully described [2–4]. In summary, the method is as simple as preparing a solution of the protein at the desired pH in the presence of a specified molar excess of the small aromatic aldehyde, freezing the solution, and lyophilizing to dryness. pH of the solution and mole ratio of aldehyde to derivatizable residues on the protein have been varied, using average degree of substitution measurements to monitor the dependence and extend of completeness of the reaction on the conditions being varied. In preparation for average degree of substitution determination, derivatized protein is dissolved

* Corresponding author. Tel.: +1 860 441 5932; fax: +1 860 441 7378.
E-mail address: thomas.r.sharp@pfizer.com (T.R. Sharp).

0731-7085/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2005.06.028

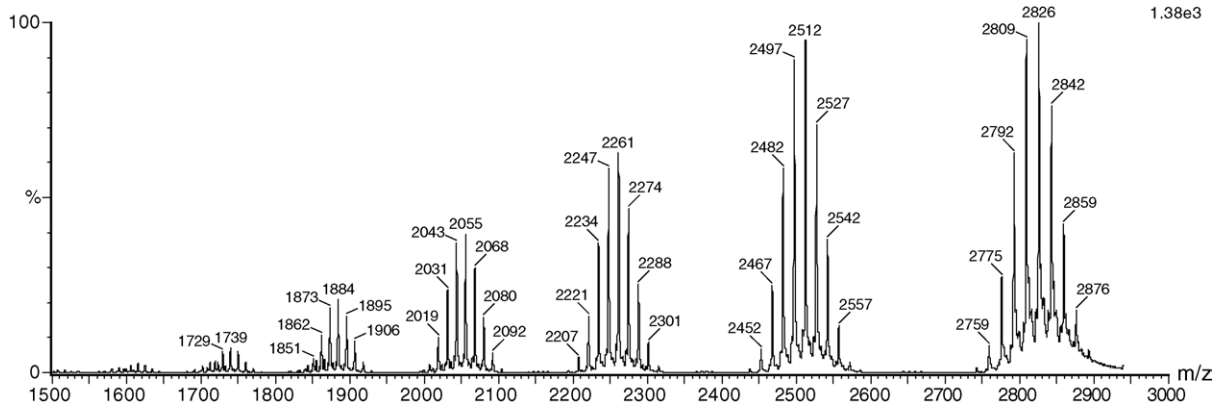


Fig. 1. Multiply charged electrospray mass spectrum of OV-pST dissolved in dry methanol. Data recorded on a Micromass QTOF-1 instrument. OV-pST was prepared from a solution starting at pH 7.

in either 2-methoxyethanol or methanol acidified with trifluoroacetic acid (TFA). The solution is continuously infused into the electrospray (ESI) ion source of the mass spectrometer with a Harvard Instruments syringe pump. The multiply charged mass spectrum is recorded and mathematically transformed to give the singly charged spectrum. Peak heights of the variously derivatized protein species are used to calculate an average degree of substitution. This measurement is used to monitor process improvement for production of the derivatized protein. Measurements have been made on a Finnigan TSQ 700 triple quadrupole system, a Micromass QTOF 1 quadrupole-time-of-flight system and a Micromass Quattro-LC triple quadrupole system.

3. Results and discussion

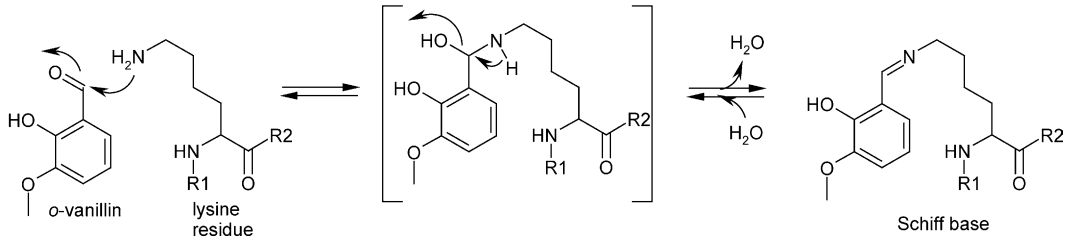
3.1. Background

The average molecular mass of pST is 21,802 Da (fully reduced). The native structure contains two disulfide bridges. The primary sequence (given above) contains 11 lysine residues. These and the amino terminus provide a potential of 12 sites for Schiff base formation between an amino group and *o*-vanillin, as shown in Scheme 1. This derivatization adds 134 Da to the relative molecular mass of the protein for each added vanillin residue.

3.2. Mass spectral observations

Fig. 1 shows a representative multiply charged ESP spectrum of ortho-vanillated porcine somatotropin (OV-pST) prepared in dry methanol. The group of peaks centered around m/z 2261 represents the +10 charge state. Computer transformation of this multiply charged spectrum, using the MaxEnt (Micromass Instruments) maximum entropy calculation procedure, produces the singly charged spectrum shown in Fig. 2. Table 1 shows an example calculation of average degree of substitution using the data from Fig. 2. The heights of peaks in the transformed spectrum (Fig. 2) are assumed proportional to the abundance of protein molecules in the sample with the given number of derivatizations (OV #). Contributions is the number of *o*-vanillin substitutions in the sample represented by this substituted protein (OV # × abundance). The sum of abundances is proportional to the total number of protein molecules (633 in this example). The sum of contributions (3723 in this example) is proportional to the total number of substitutions in the sample. Sum of contributions divided by sum of abundances equals the average degree of substitution for this sample. The measurements calculate to an aver-

AFPAMPLSSL FANAVLRAQH LHQLAADTYK EFERAYIPEG QRYSIQNAQA AFCFSETIPA PTGKDEAQQR
SDVELLRFSL LLIQSWLGPV QFLSRVFTNS LVFGTSDRVY EKLKDLEEGI QALMRELEDG SPRAGQILKQ
TYDKFDTNLR SDDALLKNYG LLSCFKKDLH KAETYLRVMK CRRFVESSCA F



Scheme 1. Reversible Schiff base formation between *o*-vanillin and a lysine residue.

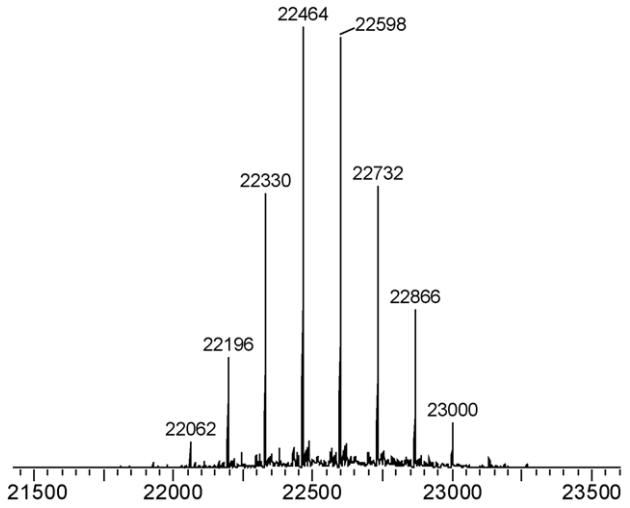


Fig. 2. MaxEnt transformed singly charged mass spectrum of OV-pST with an average degree of substitution of 5.6. Transformed from the multiply charged data of Fig. 1.

age degree of substitution of 5.6 residues/protein molecule ±1.6.

3.3. Solvent selection and quality

As the free protein is regenerated by hydrolysis of the Schiff bases, a non-aqueous solvent is necessary in which to prepare solutions for mass spectral analysis. 2-Methoxyethanol was used initially. Methanol was subsequently found to be adequate. Solvent is acidified with 0.1% (v/v) trifluoroacetic acid (TFA). Experiments with different grades of methanol showed that dry HPLC grade methanol (preferably from a freshly opened bottle) could be used to prepare solutions of OV-pST that remained stable long enough to perform the necessary mass spectral measurements. Results were reproducible from multiple determinations of average degree of substitution of the same lot of OV-pST dissolved in either 2-methoxyethanol or methanol. The same average degree of substitution result could also be obtained from a solution of OV-pST in methanol that was allowed to stand for as long as 75 min on the laboratory bench. Experiments were performed in which H_2O was purposefully added to the methanol used to prepare samples. At 0.5% H_2O by volume, a slow decline in average degree of substitution could be observed due to hydrolysis of the Schiff bases.

Table 1
Calculation of average degree of substitution

OV #	Expected mass	Abundance	Contribution
0	21800		
1	21934	2	2
2	22068	10	20
3	22202	41	123
4	22336	102	408
5	22470	164	820
6	22604	160	960
7	22738	105	735
8	22872	59	472
9	23006	17	153
10	23140	3	30
11	23274		
12	23408		

Average degree of substitution: 5.6 ± 1.6.

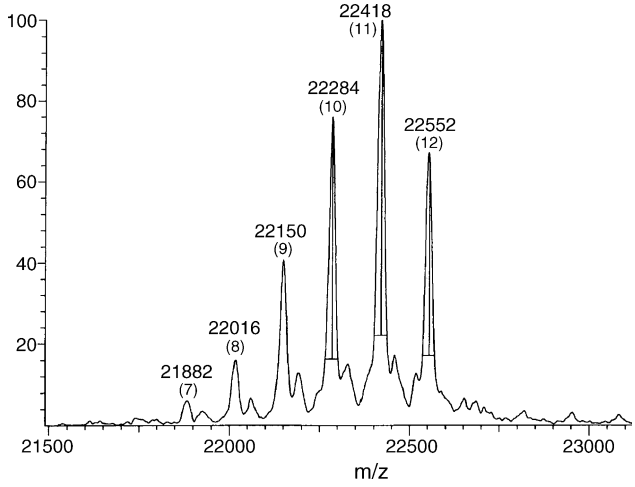


Fig. 3. Transformed singly charged mass spectrum of OV-pST with an average degree of substitution of 10.6. Data obtained on a Finnigan TSQ 700. OV-pST was prepared from a solution starting at pH 9.5. Peak tops are labeled with average masses of the species and degree of substitution in parentheses. This observation was made on a variant of the protein, with a 7 amino acid residue deletion starting at residue 31 of the sequence given in the text.

3.4. Measurement of average degree of substitution

The height of each peak in the spectrum of Fig. 2 is assumed to represent the number of protein molecules in the sample, which carry a given number of *o*-vanillin residues. If a protein molecule were drawn from this sampling, it would most likely have either five or six vanillin residues attached. The ±1.6 range represents a "standard deviation" calculated for the distribution.

This mass spectrometric method responds to differences in the average degree of substitution. The procedure for manufacturing the derivatized protein can be as simple as lyophilizing a solution of the protein and *o*-vanillin. Effects of molar ratios of protein to *o*-vanillin, of solution pH, and other manufacturing parameters on the average degree of substitution have been studied. This method provides a means of evaluating the results of those experiments.

Adjusting the pH of the solution of *o*-vanillin and protein from which the product is lyophilized promotes a different degree of substitution than from solution at neutral pH. Fig. 3 shows the transformed spectrum of OV-pST prepared at pH 9.5. Average degree of substitution calculated from the spectrum is 10.6. The fact that the distribution cuts off at a maximum of 12 *o*-vanillin residues indicates that the measurement reflects the theoretical maximum of 12 derivatizable sites on the protein, that the derivatization is specific for the lysine and n-terminal amino groups, and that the species seen

in the mass spectrum are not non-specific van der Waals clusters.

Why only five to six derivatization sites react at neutral pH, and that a larger number are accessible at higher pH, was not experimentally addressed. A crystal structure for pST is not available, but a crystal structure of human somatotropin (hST) has been deposited in the Brookhaven protein structure database [5]. Sequences of somatotropins from a total of nine species (including human and porcine) are deposited in the NCBI sequence database. Sequence comparison shows 52% sequence identity among these nine examples, and 98% positive similarity. Sequences in the NCBI database (including that for pST) still contain the "leader sequence" region. The sequence given here for our pST has had this 26 amino acid region clipped off. The protein used here was synthesized by overexpression of a gene transplanted into a bacterial system, and does not contain the leader sequence.

The six derivatization sites derivatized at neutral pH may be those exposed on the surface of the native protein and readily available for reaction. The crystal structure of hST was obtained and examined. The sites of the 11 lysine residues in pST were located in the hST structure. Sites were identified by sequence comparison. Visual assessment of exposure of the site to solvent was made. By sequence alignment with hST (and the other seven species sequences available in the NCBI database), pST shows two deletions in its overall sequence, and replacement of five of the lysine residues in pST with other residues in hST. Nevertheless, assuming the overall tertiary structure is conserved, seven lysine sites in pST are exposed to solvent. Four are buried inside the folded protein. One of the exposed sites, lysine 157 of pST, shows a 2.7 Å hydrogen bonding spatial relationship to aspartate 153, leaving 6 readily exposed lysine residues and the amino terminus. This simplistic evaluation of seven available sites corresponds to the typical average degree of substitution of six derivatizations obtained at neutral pH. Unfolding with elevated pH should expose the remaining ones. Two pairs of lysine residues in the sequence are adjacent to each other or separated by only one amino acid residue. Although steric considerations might suggest that these pairs would interfere with each other, the fact that the pH 9.5 results (Fig. 3) approach the theoretical 12 site limit – experiments have produced average degree of substitution values as high as 11.8 – suggests that this is not an influencing factor.

It is assumed that the substituted lysine residues and N-terminus of this protein have the same ionization efficiencies in the mass spectrometer as the non-substituted lysine residues and N-terminus. While equivalent ionization efficiencies have not been rigorously demonstrated, a similarity of ionization efficiencies would be expected, and empirically, Fig. 3 makes sense based on this assumption. We see in this figure a distribution of substitutions, consistent with theoretical polymer chemistry, and an abrupt end to *o*-vanillin additions at 12, which is consistent with the theoretical limit of additions in this case.

Several reports have appeared in the recent literature describing and discussing addition of small aliphatic aldehydes to proteins. 4-Hydroxy-2-nonenal is a reactive product of lipid peroxidation, capable of forming Schiff base adducts with amino side chains (e.g., lysine residues) and Michael adducts with histidine, arginine, cysteine, and lysine sidechains. Interactions with oxidized insulin B chain [6], apomyoglobin [7,8], and cytochrome c [9] have been described in detail. Derivatizations by 4-hydroxy-2-nonenal are complicated by the fact that two mechanisms are possible for derivatization with nucleophilic amino acid residues – a Michael addition across the double bond of the nonenal, and Schiff base formation with the aldehyde functional group. Differentiation of a Michael addition product versus a Schiff base is possible by the differences in the increase in mass of the two respective products. Derivatization of oxidized insulin B chain with hexanal produced Schiff base adducts only. Distribution envelopes for extensively derivatized proteins, similar to those seen here, are only shown for nonenal-derivatized apomyoglobin (between 3 and 10 adducts per protein). Stabilization of Schiff base adducts for observation required $NaCNBH_3$ reduction, in contrast to our experiences here in acidic solution without reduction.

3.5. Quantitative nature of transformations

Does MaxEnt transformation of the multiply charged spectrum to a singly charged spectrum quantitatively preserve average degree of substitution information? Visual comparison of the MaxEnt spectrum (Fig. 2) and the untransformed multiply charged spectrum (Fig. 1) suggests that MaxEnt might introduce some bias. The highest *m*/z grouping of peaks in Fig. 1, at *m*/z 2826, represents the +8 charge state. The *m*/z 2826 species corresponds to pST in the +8 charge state with six attached *o*-vanillin residues. The transformed spectrum in Fig. 2 shows the *m/z* 22,464 species as the largest peak, corresponding to protein with five attached *o*-vanillin residues, as being the most abundant peak. Average degree of substitution values were calculated from the +8 charge state for comparison. Simple peak heights, measured from the baseline, gave an average degree of substitution of 5.8. Seeing that there is a substantial baseline lift-off in this charge state distribution, correction for elevated baseline reduced the calculated value to 5.7. Comparison of the value of 5.6, determined from the transformed spectrum, suggests that the calculations taken from the transformed spectrum may underestimate the average degree of substitution slightly. While there may be a slight bias introduced in this experimental design, it should be noted that when comparing sets of data (e.g., different pH's) any bias of this type should remain constant as long as all data sets are treated the same way. The quantitative nature of transformation of multiply charged mass spectra to singly charged mass spectra, by the MaxEnt and other procedures, has been discussed elsewhere in the literature [10,11].

3.6. Mass measurement accuracy

Comparison of the mass assignments in the data set shown in Figs. 1 and 2 and Table 1 with expectations indicates a consistent deviation of 6 Da (0.03%) below the expected mass in this data set. Although greater than the 0.01% accuracy typically targeted for molecular mass measurements, these are adequate for determining average degree of substitution. This data set was chosen for illustration because it represented a worst-case example. Most of our data sets meet the 0.01% accuracy expectation. The example shown here was measured on a QTOF-1 tandem instrument. Measurements of adequate quality have been made on the QTOF-1, the Quattro-LC, and the TSQ 700 triple quadrupole instruments, showing independence of the measurement on instrument type. Indeed, a tandem instrument is not necessary.

3.7. Polydispersity of a protein sample

Synthetic polymers such as poly(ethylene glycol) are most often composed of mixtures of molecular entities. The polymerization process does not produce a single molecular entity with a precise length, but rather a mixture of molecules of varying chain length. Polydispersity is a measure of the distribution of chain lengths in the sample. The average degree of substitution measurement discussed here is a measure of polydispersity. By derivatizing pST with a variable number of *o*-vanillin residues, this polydispersity measurement, meaningless when dealing with specific proteins, becomes an important parameter to measure in characterizing the OV-pST product.

4. Conclusion

Electrospray ionization mass spectrometry is a fast, simple, and efficient way of determining ortho-vanillated substitution of lysines and the N-terminus of porcine somatotropin. This type of analysis can be performed on a single quadrupole mass spectrometer operating at fairly low mass accuracy.

Acknowledgements

Parts of this information have been previously presented [12]. We gratefully acknowledge the assistance of Dr. XiaYang Qiu in accessing the crystal structure data of hST and evaluating solvent accessibility of lysine residues in pST by analogy.

References

[1] L. Margulis, C. Speaks, S. LeMott, N. Peckham, E. Shalaev, R. Darrington, 27th International Symposium on Controlled Release of Bioactive Materials, Paris, France, 2000.
[2] M.T. Clark, R.J. Gyurick, S.K. Lewis, M.C. Murray, M.J. Raymond, United States Patent 5,198,422, 30 March 1993.
[3] T.O. Lindsey, M.T. Clark, PCT International Patent Application WO 92/01463, 6 February 1992.
[4] B.A. Hay, M.T. Clark, PCT International Patent Application WO-00/00507, 6 January 2000.
[5] L. Chantalat, N.D. Jones, F. Korber, J. Navaza, A.G. Pavlovsky, Protein Pept. Lett. 2 (1995) 333–340.
[6] F. Fenaille, P.A. Guy, J.-C. Tabet, J. Am. Soc. Mass Spectrom. 14 (2003) 215–226.
[7] M.S. Bolgar, S.J. Gaskell, Anal. Chem. 68 (1996) 2325–2330.
[8] Z. Liu, P.E. Minkler, L.M. Sayre, Chem. Res. Toxicol. 16 (2003) 901–911.
[9] A.L. Isom, S. Barnes, L. Wilson, M. Kirk, L. Coward, V. Darley-Usmar, J. Am. Soc. Mass Spectrom. 15 (2004) 1136–1147.
[10] B.B. Reinhold, V.N. Reinhold, J. Am. Soc. Mass Spectrom. 3 (1992) 207–215.
[11] J.C. Cottrell, B.N. Green, J. Skilling, 41st. ASMS Conference on Mass Spectrometry, San Francisco, CA, 31 May–4 June 1993.
[12] T.R. Sharp, R. Morris, G.J. Horan, L.H. Pezzullo, J.G. Stroh, 48th ASMS Conference on Mass Spectrometry, Long Beach, CA, 11–15 June 2000.