JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. 2003; **38**: 891–899
Published online 17 July 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jms.477

JMS

Probing of arginine residues in peptides and proteins using selective tagging and electrospray ionization mass spectrometry

Alexander Leitner and Wolfgang Lindner*

Christian Doppler Laboratory for Molecular Recognition Materials, Institute of Analytical Chemistry, University of Vienna, Währinger Strasse 38, 1090 Vienna, Austria

Received 20 January 2003; Accepted 4 March 2003

A general labelling method is presented which allows the determination of the number of guanidine groups (related to arginine and homoarginine in peptides and proteins) by means of mass spectrometry. It implies a guanidine-selective derivatization step with 2,3-butanedione and an arylboronic acid under aqueous, alkaline conditions (pH 8–10). The reaction mixture is then directly analysed by electrospray ionization mass spectrometry without further sample pretreatment. Other amino acids are not affected by this reaction although it is demonstrated that lysine side-chains may be unambiguously identified when they are converted to homoarginine prior to derivatization. Guanidine functionalities, as e.g. in the amino acid arginine, are easily identified by the characteristic mass shift between underivatized and derivatized analyte. The tagging procedure is straightforward and selective for guanidine groups. The influence of several experimental parameters, especially the pH of the solution and the choice of reagents, is examined and the method is applied to various arginine-containing peptides and to lysozyme as a representative protein. Possible applications of this technique and its limitations are discussed. Copyright © 2003 John Wiley & Sons, Ltd.

KEYWORDS: arginine recognition; butanedione; boronic acid; electrospray ionization; mass spectrometry

INTRODUCTION

In recent years, the selective recognition of amino acid residues in peptides or proteins has played an ever-increasing role in analytical chemistry, especially in combination with mass spectrometry (MS). In this respect, cysteine residues are frequently used for attaching labels for affinity purification and/or stable isotope labelling in proteomics.[1–7] Selective amino acid modification is also used to improve the mass spectrometric analysis of peptides (in addition to N- or C-terminal derivatization techniques[8]), e.g. to facilitate *de novo* sequencing or to improve detection sensitivity in general. Conversion of lysine residues to homoarginine is an example of this approach,[9,10] which can be used to increase the sensitivity of peptides with a C-terminal lysine, which suffer from a poorer ionizability in matrix-assisted laser desorption/ionization (MALDI) compared with those with arginine on the C-terminus, as shown by Krause *et al*.[11] A different lysine modification procedure for application in proteomics has also been described recently.[12]

*Correspondence to: Wolfgang Lindner, Institute of Analytical Chemistry, University of Vienna, Währinger Strasse 38, 1090 Vienna, Austria. E-mail: wolfgang.lindner@univie.ac.at
Contract/grant sponsor: Christian Doppler Society.
Contract/grant sponsor: AstraZeneca R&D Mölndal.
Contract/grant sponsor: Merck Darmstadt.
Contract/grant sponsor: piChem.

Arginine stands out from the rest of the 20 common (proteinogenic) amino acids owing to its pronounced basicity ($pK_a = 12.5$). It is involved in many important biochemical processes, e.g. as part of active sites in enzymes[13,14] and in RNA recognition domains.[15–17] Arginine-rich peptides or derivatives bearing guanidine groups have been shown to be promising drug carriers which allow the transmembrane transport of macromolecules.[18–21] In addition, guanidine and amidine moieties can be used to increase the membrane permeability of drugs.[22]

There are only relatively few mechanisms for the selective chemical recognition of arginine groups. Zenobi and co-workers used arylsulfonic acids for the *non-covalent* probing of arginine residues in peptides, proteins and small guanidines[23,24] and applied their technique to both MALDI and electrospray ionization (ESI) mass spectrometry (MS). They found that certain arylsulfonic acids, e.g. naphthalene-1,5-disulfonic acid, form specific adducts with the guanidine group whereas other basic functionalities such as lysine and the N-terminus remain unaffected. A structure for the complex has been proposed.[24]

Recently, molecular recognition of arginine in peptides using crown ether complexation in combination with ESI-MS has been reported,[25] although the applicability to probing of multiple arginine residues was limited. Schrader and co-workers have been working on the design of guanidine receptors, where non-covalent complexation is achieved with

JMS

bisphosphonate moieties,[26–30] although these interactions have not been followed with MS. Guanidinium complexation with other types of group-specific selectors has also been reported.[31,32]

Derivatization (i.e. *covalent* modification) of arginine residues in peptides using 2,4-pentanedione[33] or its trifluoromethyl-substituted analogue[34] has been employed to alleviate the problematic fragmentation behaviour of arginine-containing peptides in MALDI-MS using post-source decay. However, the most widely used derivatization reaction for arginine residues has been the reaction with α-dicarbonyl compounds such as glyoxal, phenylglyoxal or 2,3-butanedione. It has been used since the 1960s to identify the presence of arginine as part of the active site of enzymes by modifying arginine residues in biomolecules[13] when techniques such as mass spectrometry and x-ray analysis were in their infancy. Today, this approach is still frequently applied in biochemical research in order to deactivate arginine-dependent enzymes.

In this paper, we report on a modification of the 'classical' reaction of guanidine groups with α-dicarbonyl compounds. This reaction results in the formation of a dihydroxyimidazoline ring with two adjacent hydroxy groups (see Fig. 1). After the addition of 2,3-butanedione, the resulting vicinal diol is complexed with an arylboronic acid to form a bicycle comprising a boronic ester functionality. The two-step (but 'one-pot') reaction has been optimized regarding its reaction conditions and reagents using ESI-MS detection and the concept was found to be applicable to the selective labelling of peptides and proteins.

EXPERIMENTAL

Materials

All common chemicals and solvents were of analytical grade or higher purity and were obtained from Fluka (Buchs, Switzerland), Riedel-de Haën (Seelze, Germany) or Merck (Darmstadt, Germany). Water was doubly distilled prior to use.

3-Aminophenylboronic acid monohydrate, 2,3-butanedione (BD), glyoxal (40% solution in water), DL-tryptophan, DL-phosphoserine and O-methylisourea sulfate were purchased from Fluka, phenylboronic acid (PBA), phenyl-d_5-boronic acid, 1,2-cyclohexanedione and (1*S*)-(+)-camphorquinone from Aldrich (Steinheim, Germany), DL-cysteine, DL-cysteic acid, DL-histidine and lysozyme (from chicken egg-white) from Sigma (Steinheim, Germany) and furan-2-boronic acid, thiophene-3-boronic acid and pyrimidine-5-boronic acid from Frontier Scientific (Carnforth, Lancashire, UK). 4-Hydroxy-3-nitrophenylglyoxal was synthesized according to Borders *et al*.[35]

All peptides were obtained from Bachem (Heidelberg, Germany) and were used without further purification. Peptide concentrations were calculated using peptide content and purity as given by the manufacturer.

Synthesis of 3-acetamidophenylboronic acid

3-Acetamidophenylboronic acid was synthesized according to the following protocol: 155 mg (1 mmol) of 3-aminophenylboronic acid monohydrate were dissolved in 5 ml of acetonitrile and 20 µl of ammonia solution (25% in water) were added. To this solution 940 µl (10 mmol) of acetic anhydride were added in small portions with stirring. After 3 h, the solvent and the excess acetic anhydride were removed *in vacuo*. To facilitate removal of residual acetic anhydride, the residue was suspended in anhydrous THF (~2 ml) and, after evaporation, white crystals were obtained (170 mg, 0.94 mmol, 94%). The identity and purity of the product were confirmed by ESI-MS and ^{1}H NMR spectroscopy (400 MHz, solvent D_2O).

Instrumentation and mass spectrometry

A Perkin-Elmer Sciex API 365 triple quadrupole mass spectrometer (Perkin Elmer Sciex, Thornhill, ON, Canada) equipped with a pneumatically assisted electrospray ion source was used for all experiments. Samples were introduced into the instrument via a 50 µm i.d. fused-silica capillary (Composite Metal Services, Hallow, Worcestershire, UK) using a syringe pump from Harvard Apparatus (South Natick, MA, USA) at a flow-rate of 5 µl min^{-1}. Spectra were acquired in the positive ion mode by scanning over an appropriate m/z range at a scan rate of 250 Th s^{-1} and a step size of 0.1. The spectra shown in the figures represent 10–25 averaged scans. Typical instrument settings were as follows: ionization voltage 4.0 kV, declustering potential 15 V and focusing potential 140 V. For MS/MS experiments, nitrogen was used as the collision gas.

The mass spectrometer was controlled by Analyst 1.2 software (Applied Biosystems/MDS Sciex, Concord, ON, Canada). For peptide and protein mass calculations, the program massXpert[36] was used.

General labelling procedure for guanidine groups

A solution of 10 mM ammonium acetate in water, adjusted to the appropriate pH (i.e. 8–10) with ammonia solution,

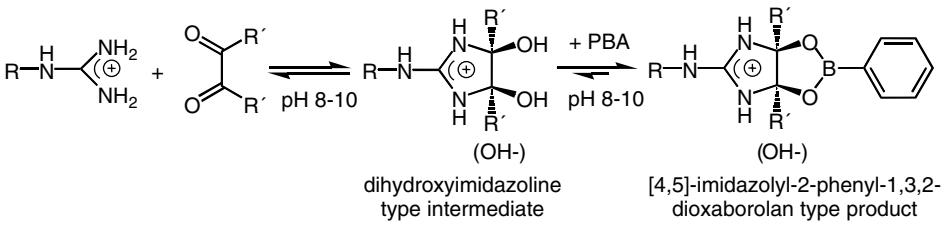

Figure 1. General derivatization scheme of guanidine groups with α-dicarbonyl compounds and phenylboronic acid, illustrating the two-step reaction. For 2,3-butanedione, $R' = CH_3$.

was added to an equal volume of a solution of the analyte in water so that the final concentration of the analyte was ~50 μmol l^{-1}. Except for the studies on the influence of reagent excess, aqueous solutions of the dicarbonyl compound (e.g. 2,3-butanedione, BD) and the arylboronic acid (e.g. phenylboronic acid, PBA) were added so that the molar excess of BD and PBA per guanidine group was ~50- and ~100-fold, respectively.

Mass spectra were either acquired by infusion of these reaction solutions via a syringe pump after various reaction time intervals (e.g. 15/60/120 min), or the progress of the reaction was monitored by continuous infusion directly after the start of the derivatization reaction.

RESULTS AND DISCUSSION

Principle of the reaction

The reaction between α-dicarbonyl compounds and guanidine compounds proceeds only relatively slowly under alkaline conditions and is reversible when diones are used. In the 1960s it was observed that the reaction proceeded significantly faster when it was performed in borate buffer.[13] This was attributed to the complexation of the diol group with the borate anion, a reaction already well documented at that time. For our strategy, we substituted the borate buffer with a solution of an arylboronic acid which should serve the same purpose, which is to shift the equilibrium of the first (condensation/cyclization) step to the product side by formation of a bicyclic 2-substituted[4,5]-imidazolyl-1,3,2-dioxaborolan-type moiety (see Fig. 1). In addition, a stable and defined complexation product for the second step should ensue and allow, in contrast to the borate-accelerated reaction, the mass spectrometric detection of the boronic ester system.

To test our hypothesis, we performed initial experiments with a model arginine-containing peptide (Lys-[Ala^3]-bradykinin) with the sequence KRPAGFSPFR. A solution of the peptide in water was mixed with a 10 mM ammonium acetate solution adjusted to pH 8.5 with ammonia solution. The system was chosen as to make direct examination of the reaction by ESI-MS possible. To the alkaline peptide solution, we added BD and PBA in 50- and 100-fold excess per arginine residue, respectively. After a reaction time of 2 h, the sample was examined by ESI-MS. The mass spectrum (data not shown) exhibited mainly the derivatization product, i.e. the addition of BD and PBA to the two arginine residues, resulting in a shift of 172 Th on the m/z scale for the doubly protonated peptide ion, $[KR^*PAGFSPFR^* + 2H]^{2+}$ (modified arginine residues are denoted R* throughout the text). With the exception of the only partially derivatized intermediate, no by-products of the reaction were observed.

Optimization of the reaction

After these encouraging preliminary results, the reaction conditions were optimized, again using the model system Lys-[Ala^3]-bradykinin + BD + PBA.

Reaction conditions

For these experiments, continuous infusion of the sample solution over a longer period of time proved to be very helpful since it allowed the monitoring of the reaction progress. The derivatization was monitored for 60–90 min and mass spectra were averaged over a 1–2 min period. Examples are shown in Fig. 2(a)–(c). The relative intensities of the signals for underivatized (m/z of $[M + 2H]^{2+} = 582.0$), partially derivatized (m/z of $[M + 2H]^{2+} = 668.0$) and completely derivatized (m/z of $[M + 2H]^{2+} = 754.0$) analytes were then plotted against reaction time as shown in Fig. 2(d) for the derivatization of Lys-[Ala^3]-bradykinin at pH 9.0. The assumption was made that the ionization efficiency of free and derivatized peptides was comparable, which was found to be reasonable for the peptides investigated, because the absolute intensities were similar for the free peptide at the beginning and the derivatized peptide at the end of the reaction. Judging from several experiments that were performed in triplicate, the variation of individual data points was in the range of 5–10%.

Not surprisingly, the solution pH has the most significant influence on the progress of the reaction, since the formation of the boronic ester is known to be very much pH dependent. An increase of one pH unit (from 7.5 to 8.5 and from 8.5 to 9.5) resulted in a 10-fold faster formation of the bicyclic reaction products. pH values above 10 were not examined in view of the possible instability of the analytes under these conditions. Other variables investigated were reaction temperature and analyte/reagent ratio. Increasing the temperature was found to accelerate the reaction only moderately and it was found that a 50-fold excess of dione and a 100-fold excess of arylboronic acid were sufficient to guarantee rapid and quantitative derivatization.

Choice of reagents

Several other α-dicarbonyl compounds were examined for their suitability for this approach. In addition to 2,3-butanedione, glyoxal, 1,2-cyclohexanedione, 4-hydroxy-3-nitrophenylglyoxal and camphorquinone were examined. More bulky reagents slowed the reaction significantly, although the proposed reaction products were observed in all cases (data not shown), while the use of glyoxal resulted in a complex reaction pattern owing to competing dehydration reactions, which make the second complexation step impossible. The tendency of glyoxal to form various reaction products with arginine has been reported previously.[37]

We evaluated several other boronic acids to compare their reactivities with that of PBA. Furan-2-boronic acid and thiophene-3-boronic acid were found to react comparably to PBA whereas derivatization with pyrimidine-5-boronic acid proceeded more slowly (data not shown).

The solubility in water is, however, a significant problem for some more hydrophobic arylboronic acids. Even PBA itself is only moderately soluble in water (maximum concentration in the range of 100 mM). Therefore, 3-acetamidophenylboronic acid (AA-PBA) was synthesized; it is easily obtained from the reaction of 3-aminophenylboronic acid and acetic anhydride (see Experimental). AA-PBA is much more soluble and exhibits a reactivity almost identical with that of PBA itself. However, it has an increased tendency to form di- and oligomers under the ESI conditions employed, which may be disadvantageous for analytes in

JMS

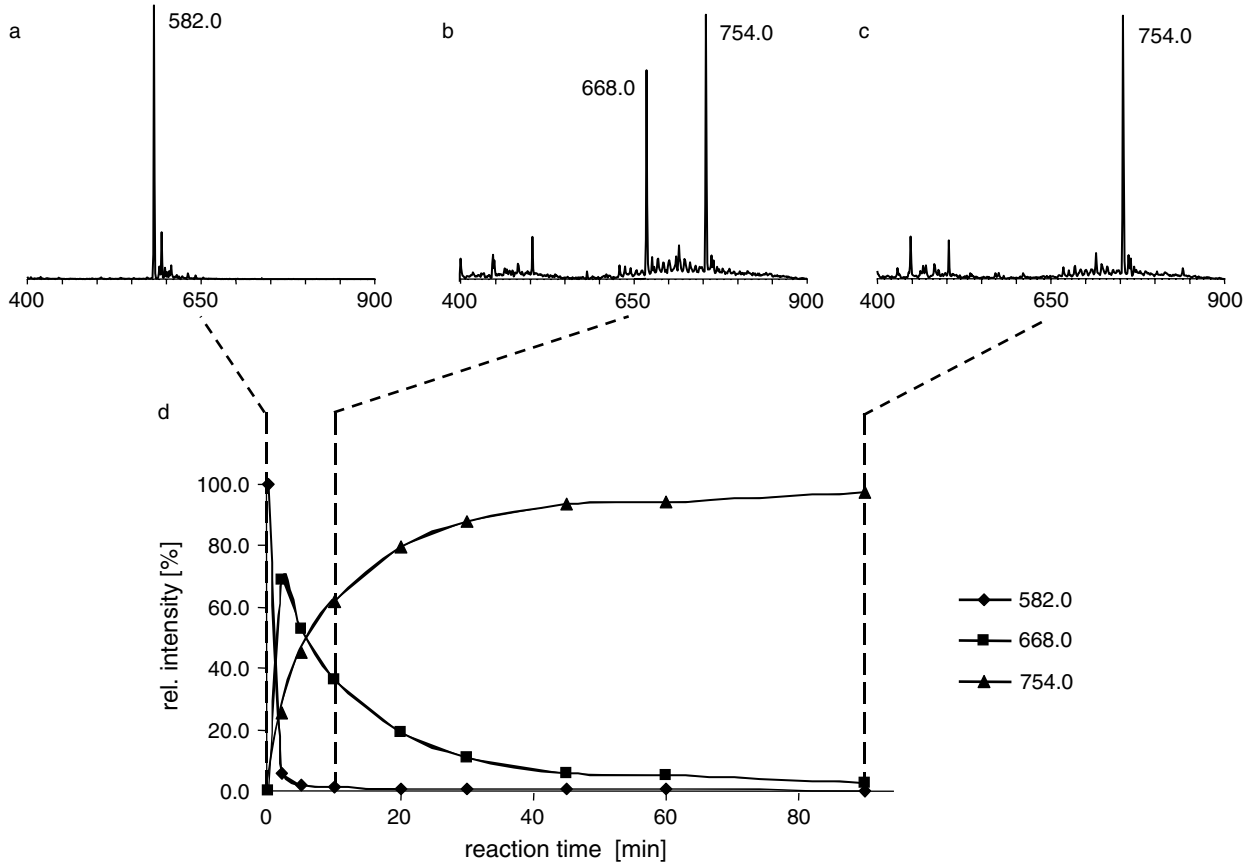

Figure 2. Modification of Lys-[Ala^3]-bradykinin (K**R**PAGFSPF**R**) with 2,3-butanedione and phenylboronic acid, pH 9.0. Spectra (a) before the addition of the reagents and (b) 10 and (c) 90 min after addition of the reagents; (d) relative intensity (%) of the signals of free (*m/z* 582.0), partially derivatized (*m/z* 668.0) and completely derivatized peptide (*m/z* 754.0) as a function of the reaction time (in minutes). For reaction conditions, see Experimental.

the lower m/z range since the clusters may overlap with the analyte signal.

ESI-MS performance

Despite the fact that ESI was performed with pure aqueous solutions, without the presence of organic solvents, the sensitivity and signal stability were completely satisfactory in routine use. Several experiments regarding the possible improvement of the signal intensity were performed. The sensitivity for Lys-[Ala^3]-bradykinin in the ammonium acetate–aqueous ammonia solution (pH 9.5) was identical (a difference of less than 5% in both intensity and signal-to-noise ratio) with that in 0.05% trifluoroacetic acid, so it can be said that the alkaline pH does not have a significant negative influence on the detection of labelled versus unlabelled arginine-containing peptides. The addition of the reagents (dione, arylboronic acid) had no detrimental effect on the sensitivity or stability of the spray, even when the reaction was continuously monitored for 90 min.

While it would be possible to add organic solvents (e.g. acetonitrile) to the sample after the derivatization step, it was found that the signal-to-noise ratio could only be improved by a factor of two through the addition of CH_3CN or CH_3OH in various concentrations (10–75%, v/v), despite the fact that the instrumental parameters were optimized for each solvent system. However, different ion source set-ups might show more pronounced effects regarding the solvent composition or the presence of the reagents.

Application to Arg-containing peptides

A variety of arginine-containing peptides were derivatized to gain an insight into possible limitations of the tagging strategy. An overview of the model peptides is given in Table 1. Arg-rich peptides were chosen to see if there is any limitation for the derivatization of repetitive R-sequences. Even peptide ε (ERMRPRKRQGSVRRRV) with an RRR motif and seven arginines in total was completely modified on all residues in less than 30 min at pH 9.5. Since the side-chains of adjacent arginine residues tend to face in opposite directions in order to minimize Coulombic repulsion of the positively charged groups, steric hindrance is not observed. For all peptides examined, no other inexplicable signals were found in the spectra which could result from modifications of other amino acid residues. Amino acids not present in the peptides (cysteine, histidine and tryptophan) and also phosphoserine and cysteic acid were added to Lys-[Ala^3]-bradykinin during the labelling step to guarantee that they are not influenced by the derivatization protocol and do not have any negative effect on the arginine modification procedure.

Preliminary tandem MS experiments were performed to investigate the fragmentation of the Arg-labelled peptides. As shown for des-Arg^9-bradykinin (Fig. 3), the tag stays

Table 1. Overview of the peptides investigated using the tagging approach with BD and PBA, with arginine residues marked in bold[a]

Sequence (common name)	M_r	p*I*	No. of arginines
Bradykinins and derivatives:			
K**R**PAGFSPF**R** (Lys-[Ala^3]-BK)	1161.6	12.01	2
RPPGFSPF (des-Arg^9-BK)	903.5	9.75	1
PPGFSPF**R** (des-Arg^1-BK)	903.5	10.18	1
RPPGFSPF**R** (BK)	1059.6	12.00	2
K**R**PPGFSPF**R** (Lys-BK)	1187.7	12.01	2
Arg-rich peptides:			
RRR	486.3	12.30	3
RK**R**TL**RR**L	1097.7	12.48	4
Q**RR**Q**R**KS**RR**TI	1483.9	12.60	5
E**R**M**R**P**R**K**R**QGSV**RRR**V	2066.2	12.40	7
Small Arg-containing peptides:			
RGD	346.2	6.09	1
RGDS	433.2	5.84	1
GP**R**	328.2	9.75	1
LWM**R**	604.3	9.75	1
GP**R**P	425.2	9.75	1
F**RR**	477.3	12.00	2
RKDV (thymopoietin II)	516.3	11.00	1
Acidic peptides:			
DSDP**R**	588.3	4.21	1
RNIAEIIKDI	1183.7	6.07	1
Lys-containing peptides:			
K**R**PPGFSPF**R** (Lys-BK)	See above		2 (1 Lys)
K**R**PAGFSPF**R** (Lys-Ala^3-BK)	See above		2 (1 Lys)
LQAAPALDKL	1038.6	5.84	0 (1 Lys)

[a] Molecular mass (M_r) is given in the monoisotopic form; p*I* values were calculated using the ExPASy p*I* tool (available at http://www.expasy.org/tools/pi_tool.html).

intact during collision-induced dissociation, yielding several b-ions (bearing the modified arginine residue) and y-ions (no modification). However, we also observed in some cases that losses of water and of the phenylboronic acid moiety can be dominant pathways. Currently, we are performing a detailed study to address the fragmentation behaviour of tagged peptides.

We also investigated if lysine side-chains can be derivatized after their conversion to homoarginine by reaction with *O*-methylisourea (OMIU). Several Lys-containing peptides were converted according to the protocol of Beardsley and Reilly,[10] which allows a very rapid modification in 10–15 min by using very high concentrations (1 g ml^{-1}) of OMIU and a temperature of 60 °C. After clean-up by solid-phase extraction, the peptides were examined by ESI-MS to verify the guanidination and then were derivatized using the usual protocol. No modifications of the N-terminal amino group of the sample peptides were observed under the conditions of the guanidination step, confirming the findings of Beardsley and Reilly,[10] who only observed some N-terminal guanidination when glycine was present on the N-terminus. The side-reactions of the procedure in complex mixtures have been addressed by the same group.[38]

The results obtained indicate that the combined modification procedure works very well. An example is given in Fig. 4, where the spectra of the peptide LQAAPALDKL (underivatized, guanidinated and guanidinated and modified) are shown. As can be seen from Fig. 4(b), after the guanidination step only a small amount of the unmodified peptide remains ($[M + 2H]^{2+}$ at m/z 520.6), while most of the analyte is present in the guanidinated form ($[M + 2H]^{2+}$ at m/z 541.6), which can be quantitatively tagged (Fig. 4(c), m/z 627.6).

Investigation of the (chemical) stability of the complexes

Boronic esters are known to be unstable at low pH, leading to the dissociation of the adduct, but the stability of the derivatives under the reaction conditions (pH 8–10) was still unknown. To elucidate their behaviour, comparative labelling studies using non-deuterated (d_0-PBA) and deuterated (d_5-PBA) phenylboronic acid were designed to reveal whether an exchange of the labels is taking place even at high pH (i.e. 9–10). Accordingly, Lys-[Ala^3]-bradykinin (containing two arginines) was derivatized with either d_0-PBA or d_5-PBA in two individual batches of the same peptide concentration, which were mixed subsequently in a 1 : 1 ratio and analysed by ESI-MS (recall that the excess of PBA is 100-fold). From the resulting spectrum, it can be seen that not two individual signals, as expected, but three are observed from the doubly protonated, doubly labelled peptide (see Fig. 5): one species incorporating two d_0-PBA esters (m/z 754.1), one incorporating two d_5-PBA esters (m/z 759.1) and an intermediate form consisting of one d_0- and of one d_5-form (m/z 756.6), obviously resulting from an exchange of the labels after mixing. From these experiments it can be deduced that the PBA exchange is completed after ~60 min at pH 9.0, resulting in a statistical 1 : 2 : 1 distribution (not accounting for isotope effects) of $2d_0$, $d_0 + d_5$ and $2d_5$ forms. Furthermore, the exchange is only slightly slowed at higher pH (data not shown).

Apparently, under the given conditions, there still exists an equilibrium of free and complexed PBA in the alkaline pH range, which does not play a role in the simple derivatization reaction, since a large excess of the reagents is used, but which makes comparative labelling studies impossible.

As a further consequence, liquid chromatographic (LC) separation of the derivatives, even under high pH conditions, would lead to a disintegration of the adduct because the excess of reagent which shifts the equilibrium is chromatographically removed. Therefore, the use of LCMS techniques for these modified analytes would not easily be possible, although it seems feasible using excess PBA as a mobile phase component.

On the other hand, the present technique easily allows the determination of the number of arginine residues present in a peptide mixture as shown in Fig. 6. A mixture of six peptides containing from no to five arginine residues (labelled A–F, see Table 2) was tagged with BD and PBA and the spectra before and after labelling were compared. In Fig. 6(b), it can be seen that after the modification step, all

JMS

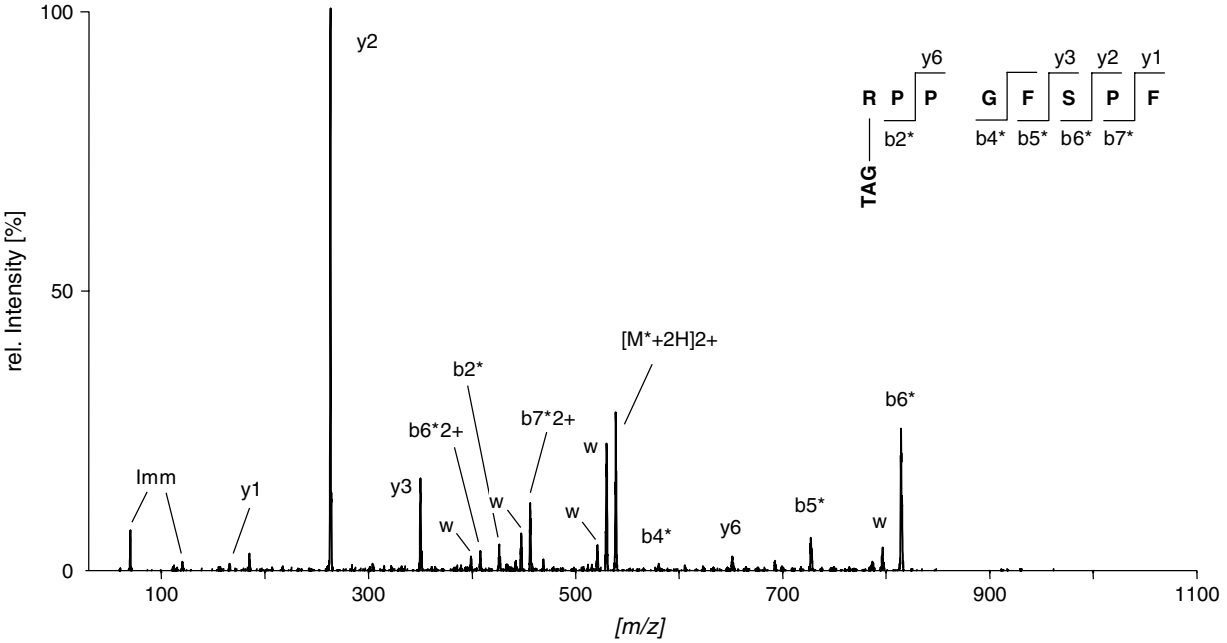

Figure 3. Tandem mass spectrum of the doubly charged precursor of labelled des-Arg^9-bradykinin (**R***PPGFSPF) recorded at a collision energy of 45 eV. Imm denotes immonium ions, w the loss of water.

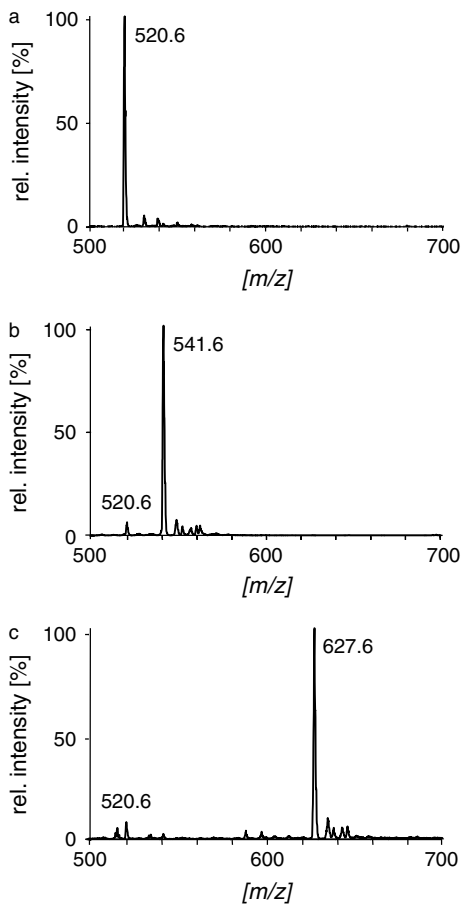

Figure 4. Example for the guanidination/homoarginine modification strategy with the Lys-containing peptide LQAAPALD**K**L. Spectra (a) before guanidination, (b) after guanidination and solid-phase clean-up and (c) after guanidination, solid-phase clean-up and modification with 2,3-butanedione and phenylboronic acid. For reaction conditions, see Experimental.

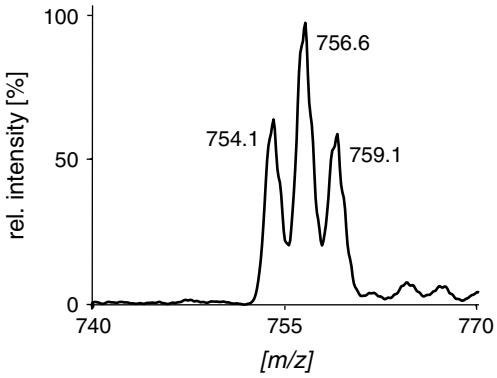

Figure 5. Exchange of the arylboronic acid part of the tag under alkaline conditions. Lys-[Ala^3]-bradykinin (K**R**PAGFSPF**R**) was derivatized in separate vials with d_0- and d_5-PBA and the batches were then mixed in a 1 : 1 ratio. The *m/z* range of the $[M^* + 2H]^{2+}$ ions (obtained 60 min after mixing) is shown. For reaction conditions, see Experimental.

signals except the one from the arginine-free peptide (B) are shifted, according to the arginine content of the peptides, by increments of 172 Da (i.e. 86 Th for $[M + 2H]^{2+}$ ions and 57.3 Th for $[M + 3H]^{3+}$ ions) when compared with the underivatized mixture in Fig. 6(a). This information, strictly related to guanidine groups, can serve as an additional, potentially valuable constraint in database searches, similar to the approach taken by Karty *et al*.[39] The mass-to-charge ratios of free and derivatized forms are given in Table 2.

Application of arginine-selective tagging to proteins

To investigate the suitability of the derivatization approach for larger polypeptides or proteins with significant tertiary structure, lysozyme was chosen as an example. Lysozyme is a 14.2 kDa protein consisting of 129 amino acids, 11 of

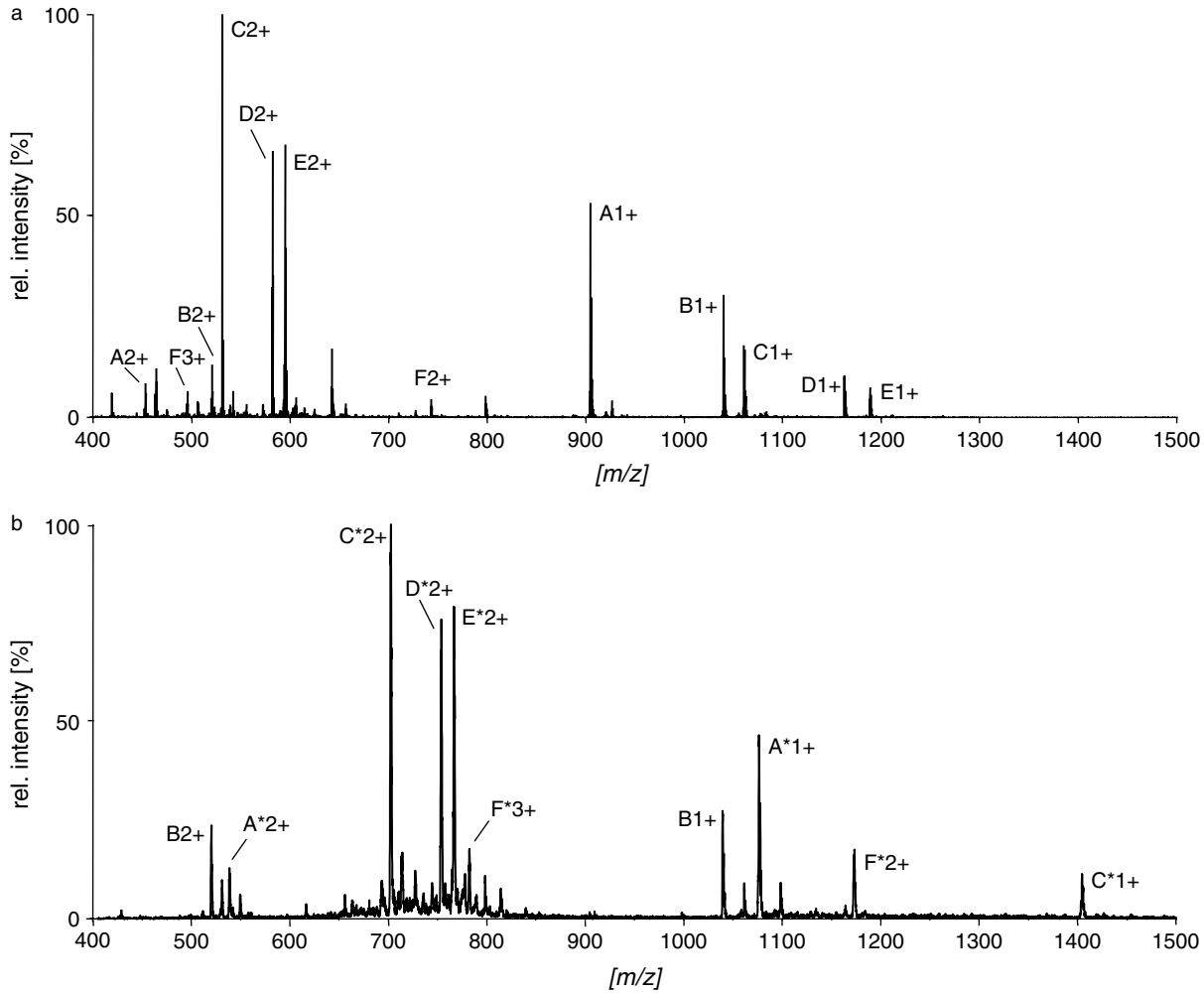

Figure 6. Arginine labelling of a mixture of six peptides (marked A–F) illustrating the suitability of the derivatization approach for peptide mixtures. Spectra are shown (a) before and (b) after modification with BD and PBA at pH 9.5 for 45 min. Peptide sequences and corresponding mass-to-charge ratios are listed in Table 2. An asterisk denotes the completely derivatized form of the peptide. For reaction conditions, see Experimental.

Table 2. Sequences (with arginine residues marked in bold) and mass-to-charge ratios of the peptides from Fig. 6[a]

ID	Sequence	n	$[M + nH]^{n+}$ theor.	$[M^* + nH]^{n+}$ theor.
A	**R**PPGFSPF	1	904.5	1076.6
		2	452.7	538.7
B	LQAAPALDKL	1	1039.6	1039.6
		2	520.3	520.3
C	**R**PPGFSPF**R**	1	1060.6	1404.8
		2	530.8	702.9
D	K**R**PAGFSPF**R**	1	1162.7	Not observed
		2	581.8	753.9
E	K**R**PPGFSPF**R**	1	1188.7	Not observed
		2	594.8	766.9
F	Q**RR**Q**R**KS**RR**TI	2	743.0	1173.2
		3	495.6	782.4

[a] An asterisk denotes the completely derivatized form of the peptide; n = charge state.

which are arginines. Lysozyme has, in its native solution conformation, several α-helical regions and a β-sheet part and has four disulfide bonds, leading to a very compact structure,[40,41] possibly leading to problems with the tagging of every arginine side-chain. However, applying the labelling approach with BD and PBA at pH 9.5, all 11 Arg residues were accessible for the modification. After a reaction time of 90 min, ~70% of the protein was modified on all Arg side-chains and 30% was modified on 10 out of 11, as shown in Fig. 7(b). No other signals stemming from protein derivatized to a lesser extent could be observed. The large shift in the molecular mass (172 per arginine) facilitates the identification of the derivatives even at high charge states without the need for high-resolution MS instrumentation.

As the charge state distribution for lysozyme in the reaction medium (at pH 9.5) is identical with that at neutral pH (data not shown), with charge states +7 to +9 dominating (+8 being the most intense signal), it may be assumed that the tertiary structure of the protein is more or less preserved. An unfolded conformation would most likely result in a shift to higher charge states since more residues able to carry the positive charge would become accessible. Owing to the relatively rigid structure of lysozyme caused by the disulfide bonds, denaturation is less likely to occur than for

JMS

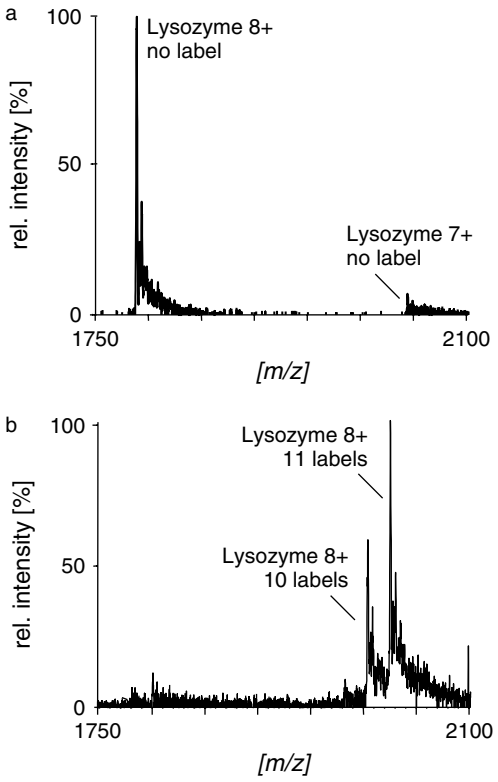

Figure 7. Labelling of arginine residues in lysozyme (~14 200 Da) using 2,3-butanedione and phenylboronic acid. Spectra are shown (a) before and (b) after the reaction at pH 9.5 for 90 min. 7+ and 8+ stand for $[M + 7H]^{7+}$ and $[M + 8H]^{8+}$, respectively.

other commonly studied model proteins, e.g. cytochrome *c*. Despite this fact, the actual conformation does not interfere with the presented derivatization concept and protocol.

Although this preliminary result proves that the approach should also be suitable for proteins, there is still limited information regarding protein conformation under the conditions employed for the labelling step and/or differences in reactivity of arginine residues depending on tertiary structure elements. This is a topic of ongoing research in our group.

CONCLUSION

A method for the selective recognition/tagging of arginine residues in peptides via the derivatization of the guanidino group with a dione and an arylboronic acid under alkaline conditions (pH 8–10) has been developed. The labelling is performed in a straightforward manner using readily available reagents under conditions specially chosen for subsequent analysis by ESI-MS without further sample handling. The labelling procedure is selective for arginine (or other analytes with guanidine functionalities) and is not disturbed by the presence of other amino acids, phosphorylation or cysteine oxidation. After guanidination of the ε-amino group, lysine residues can also be tagged.

Analysis of model peptides and peptide mixtures under the given conditions (alkaline pH, presence of reagents) revealed no negative influence on the sensitivity or stability of the ESI-MS detection. Reaction progress may be followed by continuous infusion of the sample mixture and can be used for the evaluation of modifications of the method.

Preliminary studies on the modification of Arg residues in proteins have been promising and will be the focus of future investigations, along with more detailed studies on peptides.

Acknowledgements

Financial support of the Christian Doppler Society (Vienna, Austria), AstraZeneca R&D Mölndal (Mölndal, Sweden), Merck Darmstadt (Darmstadt, Germany) and piChem (Graz, Austria) is gratefully acknowledged. The authors thank Peter Zöllner for helpful comments and critical review of the manuscript.

REFERENCES

1. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol.* 1999; **17**: 994.
2. Adamczyk M, Gebler JC, Wu J. A simple method to identify cysteine residues by isotopic labeling and ion trap mass spectrometry. *Rapid Commun. Mass Spectrom.* 1999; **13**: 1813.
3. Goodlett DR, Bruce JE, Anderson GA, Rist B, Pasa-Tolic L, Fiehn O, Smith RD, Aebersold R. Protein identification with a single accurate mass of a cysteine-containing peptide and constrained database searching. *Anal. Chem.* 2000; **72**: 1112.
4. Conrads TP, Alving K, Veenstra TD, Belov ME, Anderson GA, Anderson DJ, Lipton MS, Pasa-Tolic L, Udseth HR, Chrisler WB, Thrall BD, Smith RD. Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and ^{15}N-metabolic labeling. *Anal. Chem.* 2001; **73**: 2132.
5. Niwayama S, Kurono S, Matsumoto H. Synthesis of *d*-labeled *N*-alkylmaleimides and application to quantitative peptide analysis by isotope differential mass spectrometry. *Bioorg. Med. Chem. Lett.* 2001; **11**: 2257.
6. Sechi S. A method to identify and simultaneously determine the relative quantities of proteins isolated by gel electrophoresis. *Rapid Commun. Mass Spectrom.* 2002; **16**: 1416.
7. Zhou H, Ranish JA, Watts JD, Aebersold R. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nature Biotechnol.* 2002; **20**: 512.
8. Roth KDW, Huang Z, Sadagopan N, Watson JT. Charge derivatization of peptides for analysis by mass spectrometry. *Mass Spectrom. Rev.* 1998; **17**: 255.
9. Brancia FL, Oliver SG, Gaskell SJ. Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun. Mass Spectrom.* 2000; **14**: 2070.
10. Beardsley RL, Reilly JP. Optimization of guanidination procedures for MALDI mass mapping. *Anal. Chem.* 2002; **74**: 1884.
11. Krause E, Wenschuh H, Jungblut PR. The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal. Chem.* 1999; **71**: 4160.
12. Peters EC, Horn DM, Tully DC, Brock A. A novel multifunctional labeling reagent for enhanced protein characterization with mass spectrometry. *Rapid Commun. Mass Spectrom.* 2001; **15**: 2387.
13. Riordan JF. Arginyl residues and anion binding sites in proteins. *Mol. Cell. Biochem.* 1979; **26**: 71.
14. Borders CL Jr, Broadwater JA, Bekeny PA, Salomon JE, Lee AS, Eldridge AM, Pett VB. A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. *Protein Sci.* 1994; **3**: 541.
15. Varani G. RNA–protein intermolecular recognition. *Acc. Chem. Res.* 1997; **30**: 189.
16. Weiss MA, Narayana N. RNA recognition by arginine-rich peptide motifs. *Biopolymers* 1998; **48**: 167.

17. Famulok M, Mayer G, Blind M. Nucleic acid aptamers–from selection *in vitro* to applications *in vivo*. *Acc. Chem. Res.* 2000; **33**: 591.
18. Fawell S, Seery J, Daikh Y, Moore C, Chen LL, Pepinsky B, Barsoum J. Tat-mediated delivery of heterologous proteins into cells. *Proc. Natl. Acad. Sci. USA* 1994; **91**: 664.
19. Mitchell DJ, Kim DT, Steinman L, Fathman CG, Rothbard JB. Polyarginine enters cells more efficiently than other polycationic homopolymers. *J. Pept. Res.* 2000; **56**: 318.
20. Futaki S, Suzuki T, Ohashi W, Yagami T, Tanaka S, Ueda K, Sugiura Y. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J. Biol. Chem.* 2001; **276**: 5836.
21. Wadia JS, Dowdy SF. Protein transduction technology. *Curr. Opin. Biotechnol.* 2002; **13**: 52.
22. Witt KA, Gillespie TJ, Huber JD, Egleton RD, Davis TP. Peptide drug modifications to enhance bioavailability and blood–brain barrier permeability. *Peptides* 2001; **22**: 2329.
23. Friess SD, Zenobi R. Protein structure information from mass spectrometry? Selective titration of arginine residues by sulfonates. *J. Am. Soc. Mass Spectrom.* 2001; **12**: 810.
24. Friess SD, Daniel JM, Hartmann R, Zenobi R. Mass spectrometric noncovalent probing of amino acids in peptides and proteins. *Int. J. Mass Spectrom.* 2002; **219**: 269.
25. Julian RR, Akin M, May JA, Stoltz BM, Beauchamp JL. Molecular recognition of arginine in small peptides by supramolecular complexation with dibenzo-30-crown-10 ether. *Int. J. Mass Spectrom.* 2002; **220**: 87.
26. Schrader T. Strong binding of alkylguanidinium ions by molecular tweezers: an artificial selective arginine receptor molecule with a biomimetic recognition pattern. *Chem. Eur. J.* 1997; **3**: 1537.
27. Schrader TH. Strong binding of arylguanidinium ions by benzylic bisphosphonates–evidence for π–cation and π–π-interactions. *Tetrahedron Lett.* 1998; **39**: 517.
28. Schrader T. Chelate complexes with the P═O double bond–a new concept for molecular recognition. *J. Inclus. Phenom. Macrocycl. Chem.* 1999; **34**: 117.
29. Wehner M, Schrader T, Finocchiaro P, Failla S, Consiglio G. A chiral sensor for arginine and lysine. *Org. Lett.* 2000; **2**: 605.
30. Rensing S, Arendt M, Springer A, Grawe T, Schrader T. Optimization of a synthetic arginine receptor. Systematic tuning of noncovalent interactions. *J. Org. Chem.* 2001; **66**: 5814.
31. Bell TW, Khasanov AB, Drew MGB, Filikov A, James TL. A small-molecule guanidinium receptor: the arginine cork. *Angew. Chem., Int. Ed. Engl.* 1999; **38**: 2543.
32. Ngola SM, Kearney PC, Mecozzi S, Russell K, Dougherty DA. A selective receptor for arginine derivatives in aqueous media. Energetic consequences of salt bridges that are highly exposed to water. *J. Am. Chem. Soc.* 1999; **121**: 1192.
33. Dikler S, Kelly JW, Russell DH. Improving mass spectrometric sequencing of arginine-containing peptides by derivatization with acetylacetone. *J. Mass Spectrom.* 1997; **32**: 1337.
34. Spengler B, Luetzenkirchen F, Metzger S, Chaurand P, Kaufmann R, Jeffery W, Bartlet-Jones M, Pappin DJC. Peptide sequencing of charged derivatives by post-source decay MALDI mass spectrometry. *Int. J. Mass Spectrom. Ion Processes* 1997; **169/170**: 127.
35. Borders CL, Pearson LJ, McLaughlin AE, Gustafson ME, Vasiloff J, An FY, Morgan DJ. 4-Hydroxy-3-nitrophenylglyoxal–a chromophoric reagent for arginyl residues in proteins. *Biochim. Biophys. Acta* 1979; **568**: 491.
36. Rusconi F, Belghazi M. Desktop prediction/analysis of mass spectrometric data in proteomic projects by using massXpert. *Bioinformatics* 2002; **18**: 644.
37. Glomb MA, Lang G. Isolation and characterization of glyoxal–arginine modifications. *J. Agric. Food Chem.* 2001; **49**: 1493.
38. Karty JA, Ireland MME, Brun YV, Reilly JP. Artifacts and unassigned masses encountered in peptide mass mapping. *J. Chromatogr. B* 2002; **782**: 363.
39. Karty JA, Ireland MME, Brun YV, Reilly JP. Defining absolute confidence limits in the identification of Caulobacter proteins by peptide mass mapping. *J. Proteome Res.* 2002; **1**: 325.
40. Smith LJ, Sutcliffe MJ, Redfield C, Dobson CM. Structure of hen lysozyme in solution. *J. Mol. Biol.* 1993; **229**: 930.
41. Schwalbe H, Grimshaw SB, Spencer A, Buck M, Boyd J, Dobson CM, Redfield C, Smith LJ. A refined solution structure of hen lysozyme determined using residual dipolar coupling data. *Protein Sci.* 2001; **10**: 677.