CHROMSYMP. 2811

# Human pepsin 3b peptide map sequence analysis, genotype and hydrophobic nature

## A.T. Jones*

*Department of Clinical Chemistry, Royal Liverpool University Hospital, Liverpool L7 8XW (UK)*

## J.N. Keen

*Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT (UK)*

## N.B. Roberts

*Department of Clinical Chemistry, Royal Liverpool University Hospital, Liverpool L7 8XW (UK)*

## ABSTRACT

Peptides from a *Staphylococcus aureus* (V8) proteinase digest of human pepsin 3b have been identified by amino acid sequence analysis. Only 137 out of 326 expected residues were detected from the C and N terminal regions of the molecule. Comparison with amino acid sequences derived from nucleotide analysis of three different pepsinogen A genes, identified 2 out of 4 possible substitutions. The presence of valine at position 30 and leucine at 291 indicates that the major pepsin component of gastric juice, pepsin 3b, corresponds to pepsinogen genotype PGA-3.

Reversed-phase chromatography of native human pepsin 3b on $C_4$ (300 Å), $C_{18}$ (300 Å) or polymer (1000 Å) columns was optimal on the $C_4$ column and gradient elution with 2-propanol rather than acetonitrile. Denaturation of the protein in guanidinium hydrochloride, urea or high pH resulted in irreversible column retention. The marked hydrophobicity of denatured pepsin 3b may thus explain why the central segment of the protein was not revealed by peptide map analysis.

## INTRODUCTION

Pepsinogens belong to a multigene family of enzymes [1] that are secreted into the stomach and acid activated to form pepsins for initial hydrolysis of ingested proteins. The DNA sequences of three human pepsinogen A variants, designated PGA-3, 4 and 5 have previously been reported [2,3]. However it is still not clear which of these correspond to the active pepsins that are visualised and isolated from human gastric juice [4].

We have previously reported the development of peptide maps for the human pepsins 1, 3a, 3b, 3c and gastricsin after digestion with *Staphylococcus aureus* V8 proteinase [5]. Amino acid sequence analysis of these peptides is thus necessary to enable comparative analysis with the nucleotide derived amino acid sequence of the pepsinogen precursor proteins. Peptide fragments of human pepsin 3b, the most abundant human pepsin isoenzyme [6] secreted into gastric juice, have now been sequenced to establish its genotype by comparison with the derived pepsinogen nucleotide sequences.

The behaviour of pepsin protein on reversed-phase chromatography under different ex-

---

* Corresponding author.

perimental conditions has also been investigated as there is a paucity of information on this protein [7]. Of particular interest was the effect of denaturation on its chromatographic properties, and how this may relate to the chromatographic behaviour of pepsin's peptide fragments.

EXPERIMENTAL

*Chemicals*

Urea, guanidinium hydrochloride (GuHCl), Trizma hydrochloride (molecular biology reagents), dithiothreitol (DTT), iodoacetamide (IAA), porcine pepsin A, *Staphylococcus aureus* V8 proteinase were obtained from Sigma (Poole, UK). Ammonium acetate (AnalaR grade) was obtained from BDH (Poole, UK), trifluoroacetic acid (TFA) from Pierce and Warriner (Chester, UK), acetonitrile (HPLC-grade S) and 2-propanol (HPLC grade) from Rathburn (Walkerburn, UK).

*Pepsin purification*

Human pepsins were prepared according to the method of Peek and Roberts [4] dialysed at 4°C against 1 m$M$ HCl and freeze dried before V8 cleavage. The human pepsin 3b for reversed-phase chromatographic studies was stored at −20°C in 50 m$M$ sodium acetate buffer pH 4.1.

*Staphylococcus aureus (V8) proteinase digest*

Human pepsin 3b was digested with V8 according to the previously described method [1]. An aliquot, 500 $\mu$g, of dry enzyme was dissolved in 500 $\mu$l of 100 m$M$ ammonium acetate containing 8 $M$ urea, 50 $\mu$l of 45 m$M$ DTT was then added and the mixture incubated at 50°C for 15 min. After cooling to room temperature 50 $\mu$l of 100 m$M$ IAA were added and the solution incubated for 10 min at 20–25°C. A 1.4-ml volume of water was added followed by the addition of 16.5 $\mu$l of 1 mg/ml V8, (1:33 enzyme:substrate ratio). The digest (final pH 6.21) was incubated for 8 h at 37°C. All reactions were stopped by either freezing to −20°C or immediate chromatography.

*Reversed-phase HPLC: V8 peptide map*

Peptides, after V8 cleavage, were separated on an Exsil 300 Å C$_{18}$ 5 $\mu$m 15 × 0.46 cm reversed-phase column (Jones Chromatography, Hengoed, UK). Peptides were detected at 220 nm (UV detector SM 3000, LDC, Stone, UK) and manually collected before storage at −20°C and freeze-drying. The binary linear gradients were developed on a low-pressure ternary mixing pump (Model CM 4000, LDC) and the chromatograms recorded on a CI 4000 computing integrator (LDC).

*Reversed-phase HPLC: Human pepsin 3b profiles*

Purified human pepsin 3b was concentrated to 26 mg/ml in a Centricon-10, 10 000 molecular mass cut-off ultrafiltration unit (Amicon, Stonehouse, UK). A 2-$\mu$l volume of pepsin solution was mixed with 50 $\mu$l of water before manual injection and HPLC gradient initiation. The elution of pepsin 3b was studied on Exsil 300 Å C$_{18}$ 5 $\mu$m, Apex wp butyl 7 $\mu$m C$_4$ (Jones Chromatography) and Polymer PLRP-S 1000 Å 8 $\mu$m (Polymer Labs., Church Stretton, UK) 15 × 0.46 cm columns.

*Denaturation experiments*

Human pepsin 3b (50 $\mu$l of 0.75 mg/ml in 50 m$M$ sodium acetate buffer pH 4.1) was added to 50 $\mu$l of 6 $M$ GuHCl in 100 m$M$ Tris–HCl, 50 m$M$ DTT (final pH 4.61), incubated at 25°C for 40 min and injected on to the C$_4$ column. A sample (50 $\mu$l) of native pepsin 3b was initially used as a control.

*Solid-phase amino acid sequence analysis*

The freeze-dried peptide was dissolved in 30 $\mu$l 50% (v/v) aqueous acetonitrile and a 10-$\mu$l aliquot was dried onto a Sequelon–arylamine (AA) membrane disc at 56°C. The disc was then cooled to room temperature (22°C) and 10 $\mu$l 10 mg/ml 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide in 2-(N-morpholino)ethanesulphonic acid (MES) buffer pH 5.0 dried onto it (about 30 min) to covalently attach the peptide (via C-terminus and acidic side chains). The AA-coupled peptide was then subjected to automated solid-phase Edman degradation on a MilliGen/

Biosearch 6600 ProSequencer (Millipore–Waters, Watford, UK). The phenylthiohydantoin–amino acids were identified (at 269 nm) by reversed-phase HPLC (Waters SequeTag $C_{18}$ column) on a Waters 600 system, using a gradient of acetonitrile in 30 m$M$ ammonium acetate, pH 4.8.

## RESULTS

The peptide map for human pepsin 3b after V8 digestion is shown in Fig. 1. Incubation of V8, without pepsin protein, indicated that the doublet peak 2 was a V8 self-cleavage product and increased with increasing time of incubation in all the V8 digests performed. No other peptide peaks were noted in the absence of pepsin 3b, thus peptides labelled 3–14 are associated with pepsin 3b after cleavage by V8. No further peaks were distinguished from baseline shifts by increasing solvent B to 100% or by changing solvent B to 0.05% (v/v) trifluoroacetic acid (TFA) in acetonitrile.

The individual peptide sequence data (Table I) in comparison with the predicted intact protein amino acid sequence (Fig. 2) indicate that V8 cleavage occurred predominantly in the C
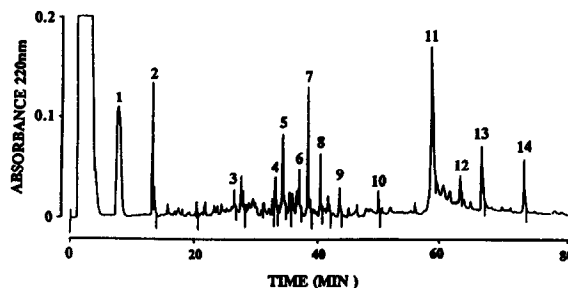


Fig. 1. Reversed-phase peptide map of human pepsin 3b digested with V8. Injection volume 1 ml (corresponding to 174 $\mu$g of digested pepsin); flow-rate, 1.5 ml/min:operating back pressure 1400–2000 p.s.i. (9653–13 870 kPa). Solvent A = 0.1% trifluoroacetic acid–water (0.1:99.9, v/v); B = solvent A–acetonitrile (40:60, v/v). Linear gradient: 100% A for 5 min, 0–30% B over 20 min and 30–80% B generated over 50 min. Peaks: 1 = DTT; 2 = V8 self cleavage products; 3–14 = human pepsin 3b peptides.

and N terminal regions of the protein. The peptides observed were in groups of overlapping residues between 1–58, 108–125, 172–187 and 280–326. Chromatographic retention time ($t_R$) generally increased with increasing chain length (Table I). However, the short sequence of peptide 10 but with relatively long retention time (42.65 min), may be explained by the poor yield (Fig. 1) and by the fact that the sequence

## TABLE I

AMINO ACID SEQUENCES OF 12 PEPTIDES COLLECTED FOLLOWING REVERSED-PHASE HPLC OF A V8 DIGEST OF HUMAN PEPSIN 3b (FIG. 1)

X indicates unidentified residues. Cysteine was determined as IAA-modified cysteine (carbamidoethyl cysteine), which runs in an identical position to glycine. Comparison with the gene sequence allowed discrimination of these two alternatives (Fig. 2).

| Peak No. | Determined peptide sequence and retention time (min) |
|---|---|
| 1 | Dithiothreitol (8.32) |
| 2 | V8 self cleavage products (13.57) |
| 3 | RANNQVGLAPVA (26.01) |
| 4 | SSYYTGSL (32.08) |
| 5 | NWVPVTVE (32.31) |
| 6 | GSCISGFQGMNLPTESGE (33.43) |
| 7 | XVPSVYXSSLAXTNHNXFNP (36.11) |
| 8 | VDEQPLENYLDME (37.46) |
| 9 | TGSSNLWVPSVYXSSLAXTNXNXFNP (39.51) |
| 10 | YYAPFDGILG (42.65). End not reached. Low yield |
| 11 | YFGTIGIGTPAQDFTVVFDTGSSNLWVPSVYCS (57.48) |
| 12 | XXPGSFLYYAPFDXXLXLAY (62.05) |
| 13 | LWILGDVFIRQYFXVFDRANNQVGLAPVA (65.45) |
| 14 | GSXISGFQGMNLPTESGELXILGDVFIXQYFTVFD (72.26) |

```
1      10        20        30        40        50        60
VDEQPLENYLDMEYFGTIGIGTPAQDFTVVFDTGSSNLWVPSVYCSSLACTNHNRFNPEDSST
VDEQPLENYLDME (37.46)                  XVPSVYCSSLACTNHNXFNP (36.11)
                         TGSSNLWVPSVYCSSLACTNXNXFNP (39.51)
           YFGTIGIGTPAQDFTVVFDTGSSNLWVPSVYCS (57.48)
```

```
    70      80        90       100       110       120
YQSTSETVSITYGTGSMTGILGYDTVQVGGISDTNQIFGLSETEPGSFLYYAPFDGILGLAY
                                 YYAPFDGILG (42.65)
                     (62.05) XXPGSFLYYAPFDXXLXLAY
```

```
   130       140       150       160       170       180
PSISSSGATPVFDNIWNQGLVSQDLSFVYLSADDQSGSVVIFGGIDSSYYTGSLNWVPVTVE
                                        SSYYTGSL (32.08)
                                         (32.31)NWVPVTVE
```

```
   190       200       210       220       230       240
GYWQITVDSITMDGEAIACAEGCQAIVDTGTSLLTGPTSPIADIQSDIGASEDSDGDMVVSC
```

```
   250       260       270       280       290       300       310
SAISSLPDIVFTIDGVQYPVPPSAYILQSEGSCISGFQGMNLPTESGELWILGDVFIRQYFT
                                GSCISGFQGMNLPTESGE (33.43)
                                          LWILGDVFIRQYFX ---
                         GSCISGFQGMNLPTESGELXILGDVFIXQYFT ---
```

```
       320
VFDRANNQVGLAPVA
     RANNQVGLAPVA (26.01)
VFDRANNQVGLAPVA (65.45)
VFD (72.26)
```
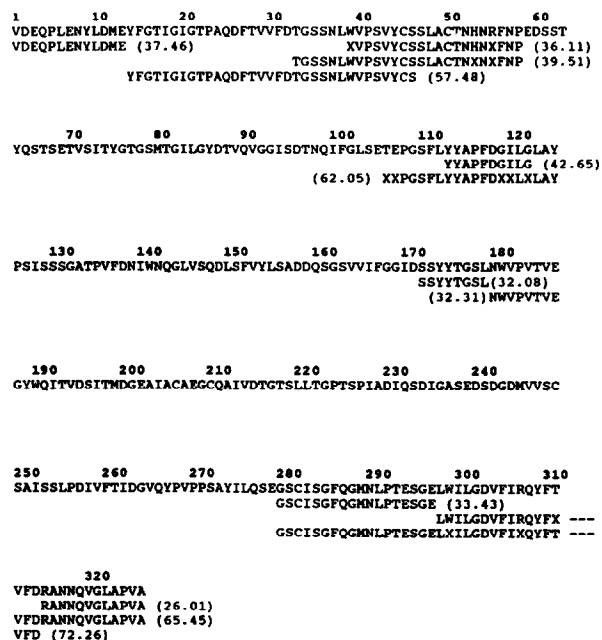
Fig. 2. Sequenced peptides and retention times from V8 digest of human pepsin 3b (Table I, Fig. 1) and the predicted amino acid sequence from human pepsinogen (PGA-3) gene analysis [2]. X is an unknown amino acid; ( ) represents the peptide chromatographic retention time (min), see Fig. 1.



Fig. 3. Reversed-phase HPLC of human pepsin 3b. Injection volume 52 $\mu$l, corresponding to 52 $\mu$g of pepsin; flow-rate, 1.0 ml/min. Linear gradient for all chromatographs: 0–85%B in 20 min. Solvent A = 0.05% TFA in water, B = 0.05% TFA in 2-propanol for Fig. 3 top (a, b and c); B = 0.05% TFA in acetonitrile for Fig. 3 bottom (d, e and f). Columns (a, d) $C_{18}$, (b, e) $C_4$ and (c, f) polymer.

analysis of the peptide was incomplete. Similarly the low yield of peptide 12 ($t_R$ 62.05 min) also suggests an incomplete sequence.

Reversed-phase chromatography of the native human pepsin 3b gave a single well-defined peak using $C_{18}$, $C_4$ or polymer-based columns (Fig. 3). Pepsin recoveries were higher with 2-propanol elution (Fig. 3 top), and in the column order: polymer > $C_4$ > $C_{18}$; however, the retention times were similar. The retention times with acetonitrile (Fig. 3 bottom) were more column dependent, being similar for the $C_4$ and polymer columns but eluted later on the $C_{18}$ column and required up to 65% (v/v) acetonitrile. Pepsin recoveries with acetonitrile elution were in the column order $C_4$ > $C_{18}$ > polymer.

Denaturation of human pepsin 3b after pretreatment with guanidinium hydrochloride, DTT and IAA (final pH 4.61) resulted in complete retention [using up to 93% (v/v) acetonitrile and 0.05% TFA in water] of the molecule on the $C_4$ column (compare Fig. 4b with a). A similar
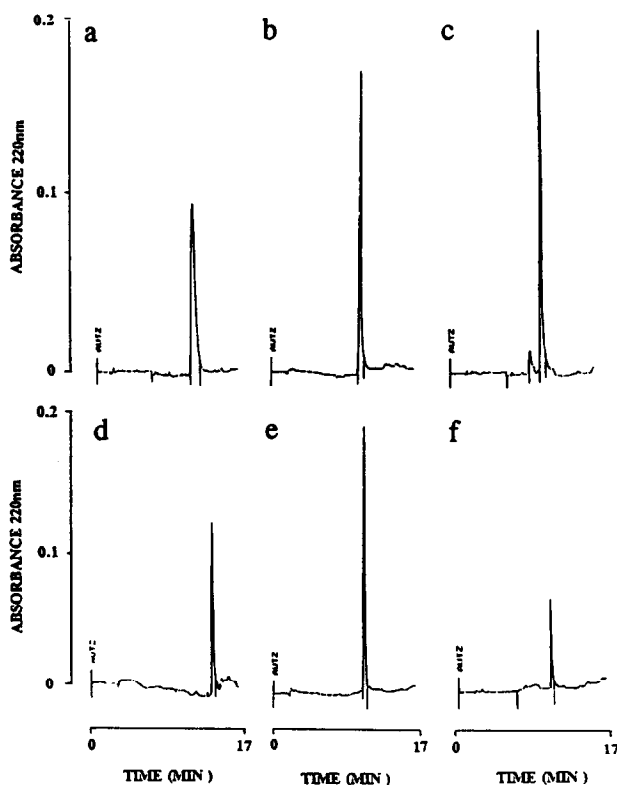
effect (results not shown) was seen on $C_{18}$ and the polymer columns and after pepsin denaturation by alkali (pH 10.0) or 8 $M$ urea.

DISCUSSION

Nucleotide sequence analysis of three human pepsinogen A variants [3], PGA-3, 4 and 5, indicates that potential substitution positions in the active enzymes are at the N and C terminal residues 30 and 291, and at central domain positions 160 and 203 (Table II). The peptide map sequence data for human pepsin 3b identified two of the four possible substitutions, namely valine at 30 and leucine at 291 suggesting therefore that pepsin 3b is the PGA-3 genotype
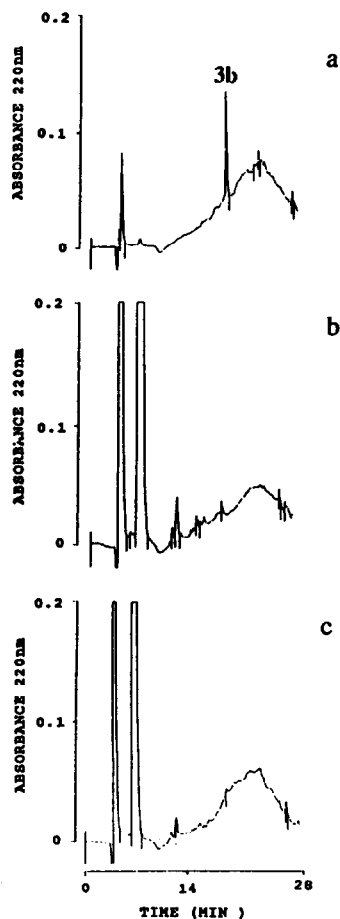
Fig. 4. Reversed-phase HPLC of human pepsin 3b before and after denaturation in 6 M guanidinium hydrochloride, 100 mM Tris–HCl, 50 mM DTT, on the C₄ column. Solvent A = 0.05% TFA, B = 0.01% TFA in acetonitrile, flow-rate 1.0 ml/min. Gradient 0–93%B in 28 min. (a) 50 μl of native 3b 0.75 mg/ml in sodium acetate buffer pH 4.1, (b) 40 μl denaturant + 50 μl pepsin, (c) 50 μl of denaturant.

TABLE II

PREDICTED AMINO ACID SUBSTITUTIONS OF PEPSINS A (PA)-3, 4 AND 5 [2,3]

| Pepsin variant | Residue position and identification | | | |
|---|---|---|---|---|
| | 30 | 160 | 203 | 291 |
| PA-3 | Val | Gln | Ala | Leu |
| PA-4 | Leu | Lys | Thr | Val |
| PA-5 | Leu | Gln | Ala | Leu |

[3]. Amino acid sequence analysis of all the peptide fragments after V8 digestion of human pepsin 3b revealed only 137 out of 326 residues or 42% of the denatured protein. The N and C terminal regions were clearly susceptible to V8 cleavage but not the central section of the molecule, which contains the other two substitution sites. The absence of this section of the molecule on peptide map sequence analysis, could have been as a result of a reduction in, or a modification of, V8 activity in the presence of urea (2 M), DTT (1 mM) and IAA (2.5 mM). 4 M urea has previously been shown to decrease V8 activity by 50% [8]. This explanation seems unlikely in view of the extensive cleavages at the N and C termini. A more plausible alternative is that the portion of the molecule not accounted for, includes the more hydrophobic fragments which are retained on the column and not eluted even with 100% acetonitrile. The several small unlabelled peaks observed on chromatography were not analysed but it is unlikely that these will contain the remaining 58% of the pepsin molecule.

The established sites of cleavage of V8 proteinase are predominantly at aspartic and glutamic acid residues [9]; however, V8 also cleaved pepsin 3b at leucine and serine residues (peptides 4, 5 and 11). The low yields of peptides 10 and 12 suggest that the C-terminus was not reached during sequencing, thus explaining the unusual cleavages at the C-terminus of glycine and tyrosine. An apparent lack of specificity of V8 cleavage has previously been observed with other proteins including two aspartic proteinases, bovine chymosin [10], and rhizopuspepsin [11]. Interestingly, 82% of the rhizopuspepsin molecule was identified by a combination of gel filtration and reversed-phase chromatography following a V8 digest. V8 digests of cyanogen bromide fragments of porcine pepsin [12,13] did however show cleavages at glutamic and aspartic acid residues only. These experiments thus emphasise the possible varied proteolytic specificities of V8 proteinase under different digestion conditions. The possibility of predicting chain length by peptide hydrophobicity and chromatographic retention time was promising but the problems of incomplete se-

quencing made it difficult to confirm a previously established relationship using synthetic peptides of known chain length [14,15].

The experiments on the chromatographic properties of native and denatured intact human pepsin 3b confirmed the marked hydrophobicity of this protein. Native pepsin was eluted from reversed-phase columns with approximately either 55% acetonitrile or 2-propanol, which is higher than that required to elute (from similar columns) the most hydrophobic protein, amyloglucosidase ($M_r$ 97 000) in a study of 33 commercially available proteins [7]. Denatured pepsin, however, was not eluted from the three selected columns even with 85% acetonitrile or 2-propanol. This may be explained by exposure to the reversed-phase material of interior hydrophobic residues due to loss of secondary protein structure. In the native state the hydrophobic core is unlikely to be exposed to the stationary phase. The column life was also significantly reduced by injecting denatured pepsin, to 2 or 3 injections (100 $\mu$g of protein) and therefore restricting possible modifications of eluting conditions. In contrast to these effects on human pepsin 3b, sample pre-treatment by addition of denaturing agents (GuHCl and urea) was reported to increase, and reducing agents (DTT) decrease respectively, the reversed-phase chromatographic recoveries of other hydrophobic proteins [16]. However, reduction of these proteins and subsequent complete retention on column was not reported.

In conclusion reversed-phase peptide mapping followed by amino acid sequence analysis has allowed the genetic identification of human pep-

sin 3b. By following the same procedure, the other pepsin isoenzymes 3a, c and 1 can similarly be assigned. Pepsin has marked hydrophobic properties under standard reversed-phase chromatographic conditions which influence recoveries of the denatured molecule and certain peptide fragments.

REFERENCES

1 R.T. Taggart, T.K. Mohandas, T.B. Shows and G.I. Bell, Proc. Natl. Acad. Sci. U.S.A., 82 (1985) 6240.
2 K. Sogawa, Y. Fujii-Kuriyama, Y. Mizukami, Y. Ichihara and K. Takahashi, J. Biol. Chem., 258 (1983) 5306.
3 M.P.J. Evers, B. Zelle, J.P. Bebelman et al., Genomics, 4 (1989) 232.
4 K. Peek and N.B. Roberts, J. Chromatogr., 476 (1989) 291.
5 A.T. Jones and N.B. Roberts, J. Chromatogr., 599 (1992) 179.
6 A.T. Jones, K.K. Balan, S.A. Jenkins, R. Sutton, M. Critchley and N.B. Roberts, J. Clin. Pathol., (1993) in press.
7 W.G. Burton, K.D. Nugent, T.K. Slattery, B.T. Summers and L.R. Snyder, J. Chromatogr., 443 (1988) 363.
8 G.R. Drapeau, Methods Enzymol., 47 (1977) 189.
9 J. Houmard and G.R. Drapeau, Proc. Natl. Acad. Sci. U.S.A., 69 (1972) 3506.
10 B. Foltman, V.B. Pedersen, D. Kauffman and G. Wybrandt, J. Biol. Chem., 254 (1979) 8447.
11 K. Takahashi, J. Biol. Chem., 262 (1987) 1468.
12 J. Marciniszyn, Jr., P. Sepulveda, W.-Y. Hunag and J.P. Lanier, J. Biol. Chem., 250 (1975) 5076.
13 P. Sepulveda, J. Marciniszyn, Jr., D. Liu and J. Tang, J. Biol. Chem., 250 (1975) 5082.
14 J.L. Meek, Proc. Natl. Acad. Sci. U.S.A., 77 (1980) 1632.
15 C. Chabanet and M. Yvon, J. Chromatogr., 599 (1992) 211.
16 K.D. Nugent, W.G. Burton, T.K. Slattery and B.F. Johnson, J. Chromatogr., 443 (1988) 381.