Structure analysis of purified histone H5 and of H5 in nuclei by limited proteolysis

Marianne HALLUPP,* Friedrich BUCK† and Wolf H. STRÄTLING*‡

*Institut für Physiologische Chemie und †Institut für Zellbiochemie und klinische Neurobiologie, Universitäts-Krankenhaus Eppendorf, Martinistrasse 52, 2000 Hamburg 20, Federal Republic of Germany

The structure of purified histone H5 in 1 M-NaClO₄ and of H5 in nuclei was analysed by digestion with either one of three endoproteinases, papain, subtilisin or elastase, which preferentially cleave unstructured protein regions (and additionally with trypsin). Digestion with papain and subtilisin produced 'limiting' resistant peptides (p1 and s1) that contain the central region between residues 18–20 and residue 114. Digestion of purified H5 with elastase generated resistant peptides e1 and e2, and that of H5 in nuclei, peptide e2. Peptides e1 and e2 contain the region from residues 22 to 114 and 109 respectively. These results show that a central region of H5 encompassing the sequence between residues 18–22 and residue 114 is folded into a compact structure. A central structured 'core' domain ranging from residues 22 to 100 is defined by the limit trypsin peptide t3, which is identical to the previously described fragment GH5 [Aviles *et al.* (1978) Eur. J. Biochem. 88, 363–371]. Generation of peptides e2 and t3, as well as of resistant peptides of lower abundance, shows that the sites near Lys-100 and Lys-109 exhibit some proteolytic sensitivity, which may result either from an exposed location or from a locally less compact conformation. Significantly, all these structural features of H5 are manifested in the purified form as well as in nuclei. A role of the structured region from residues 101 to 114 for the interaction with linker DNA and the determination of its path is discussed.

INTRODUCTION

Histone H5 and histones H1 are a distinct class of lysine-rich chromosomal proteins, which have in common their location at the periphery of the nucleosome, where they stabilize higher-order chromatin structure [1]. Differences in transcriptional activity of eukaryotic genes are frequently associated with changes in the composition of these histones. An extreme example is the replacement of most of histone H1 with histone H5 during erythropoiesis in birds, reptiles, amphibians and fish [2]. In order to understand the mechanisms underlying the participation of histones H5 and H1 in chromatin folding and in modulation of gene activity, their secondary and tertiary structure is of great interest.

Significant progress in this area was achieved by limited proteolysis in structuring ionic milieus. Treatment of isolated histone H5 with trypsin in 0.2 M-K₂SO₄ resulted in a very rapid initial digestion and the formation of a 'limiting' resistant peptide, which comprised the central sequence from residue 22 to residue 100 [3]. A broad range of physical measurements (spectroscopic, hydrodynamic, neutron-scattering and n.m.r.) has shown that this peptide has a compact globular three-helix structure [3–5]. Initially, it was concluded that the remaining segments of the molecule (*N*-terminal residues 1–21 and *C*-terminal residues 101–189) exist as 'disordered' chains in medium-ionic-strength solution [3]. Later, secondary-structure predictions suggested a high potential for α -helix formation over a great portion of the *C*-terminal chain [6]. This high potential has been shown by c.d. measurements to be realized *in vitro* under conditions known to stabilize α -helices [6]. Further, the *C*-terminal segment of H5 contains five copies of the sequence motif Ser-Pro-Lys/Gly/Ala-Lys/Arg, which has the potential to form β -turn structures [7]. Notably, the structure of H5 *in situ* has yet to be established; in the past this was mainly hampered by the inapplicability of most physical methods, such as c.d., to intact nuclei.

We have investigated the structure of purified H5 and H5 *in situ* by digestion with three endoproteinases (papain, subtilisin or elastase), which preferentially cleave unstructured protein regions. The 'limiting' resistant peptides generated show that a central region of H5, encompassing the sequence between residues 18–22 and residue 114, is folded into a compact structure. Our results also indicate that the sites near Lys-100 and Lys-109 exhibit some proteolytic sensitivity. These structural features of H5 are manifested in the purified form as well as in nuclei.

EXPERIMENTAL

Preparation of histone H5

Histone H5 was isolated as described [8]. Briefly, nuclei ($\approx 7800 A_{260}$ units) were isolated from mature chicken erythrocytes in 1 $\times$ SSC (0.15 M-NaCl/0.015 M-sodium citrate)/0.25 % Triton X-100/0.25 % SDS/1 mM-phenylmethanesulphonyl fluoride (PMSF), washed in 0.35 M-NaCl/10 mM-Tris/HCl (pH 7.5)/0.25 mM-PMSF, and extracted for 17 h at 4 °C with 0.75 M-NaCl in the same buffer. The extract was separated from the chromatin jelly by centrifugation and dialysed against column buffer [10 mM-sodium phosphate (pH 7.0)/0.25 mM-PMSF]. H1 and H5 were fractionated by ion-exchange chromatography on Whatman CM52 (1.5 cm $\times$ 5 cm column), using a linear gradient of 0.3 to 0.8 M-NaCl in column buffer. Protein fractions were monitored at 230 nm and subjected to SDS/15 %-(w/v)-PAGE. The pool of pure H5 was dialysed into 1 mM-sodium phosphate (pH 7.5)/0.2 mM-EDTA/0.25 mM-PMSF, frozen in small portions and stored at –80 °C.

Preparation of erythrocyte nuclei

Chicken erythrocyte nuclei were freshly prepared for each digestion experiment. A 2 ml portion of blood was collected in 1 $\times$ SSC and centrifuged at 3000 rev./min (1500 g). The pelleted erythrocytes were suspended in 1 $\times$ SSC/0.25 % Triton X-100 and incubated for 20 min at 4 °C, and the suspension was

Abbreviation used: PMSF, phenylmethanesulphonyl fluoride.

‡ To whom correspondence should be sent.

then centrifuged at 2000 rev./min (650 g). After repetition of this step, nuclei were washed in nuclear buffer containing 120 mM-NaCl, 3 mM-MgCl₂ and 30 mM-Tris/HCl at the pH specified for each proteolytic enzyme.

Limited proteolysis of purified H5 and H5 in nuclei

The proteolytic enzymes used were tosylphenylalanylchloromethane ('TPCK')-treated bovine pancreatic trypsin (Sigma; type XIII), papain (Boehringer Mannheim), subtilisin (Boehringer Mannheim), and porcine pancreas elastase (Sigma; type III). Papain was activated in 10 mM-EDTA/60 mM-2-mercaptoethanol/50 mM-cysteine hydrochloride, pH 3.0, for

30 min at 4 °C immediately before use. To digest purified H5, reaction conditions were as follows: 10 mM-Tris/HCl, pH 8.0, 25 °C for trypsin; 50 mM-Tris/HCl/0.1 mM-EDTA/1 mM-cysteine hydrochloride, pH 6.7, 25 °C for papain; 10 mM-Tris/HCl, pH 8.0, 37 °C for subtilisin; 10 mM-Tris/HCl, pH 8.8, 37 °C for elastase. In a parallel set of reactions, the digestion buffers additionally contained 1 M-NaClO₄. Histone H5 was mixed with twice-concentrated digestion buffer, and then incubated with each proteinase at the indicated concentrations. At the times indicated, aliquots of the proteolytic reaction mixture were removed and added to freshly prepared 100% (w/v) trichloroacetic acid (to give a final trichloroacetic acid concentration of 25%). After 1 h at 0 °C, the precipitated proteins were recovered by centrifugation, washed twice with ice-cold acetone, and dried. In a third set of reactions, purified H5 was incubated in 30% 2,2,2-trifluoroethanol/50 mM-Tris/HCl, pH 8.0, for 5 min at the enzyme concentrations indicated in Figs. 1 and 3 (below).

To digest erythrocyte nuclei, reaction conditions were as follows: 30 mM-Tris/HCl (pH 8.0)/120 mM-NaCl/3 mM-MgCl₂,

Fig. 1. Time course of digestion of purified H5 with trypsin monitored by SDS/PAGE

Purified H5 was incubated with tosylphenylalanylchloromethane-treated trypsin (4 μg/ml) at 25 °C in 10 mM-Tris/HCl, pH 8.0 (lanes 1-6; '10 mM-Tris') or in 1 M-NaClO₄/10 mM-Tris/HCl, pH 8.0 (lanes 7-11; '1 M-NaClO₄'). At the indicated times (in min and set at 90°), aliquots were removed, the proteolytic reaction terminated, and the product peptides analysed by SDS/15%-PAGE. In lanes 12-14 ('TFE'), H5 was digested with 150, 400 and 800 μg of trypsin/ml respectively for 5 min in 30% (v/v) trifluoroethanol/50 mM-Tris/HCl, pH 8.0. The arrowhead indicates the position of the trypsin-resistant central folded fragment of H5 (here named t3 and previously termed GH5; [3]). The resistant peptides t2 and t3 are most clearly visible in lanes 10 and 11. The marker lanes (H) contain total chicken erythrocyte histones extracted with 0.2 M-H₂SO₄.

Fig. 2. Time course of digestion of purified H5 with papain

Purified H5 was incubated with activated papain (30 μg/ml) at 25 °C in 50 mM-Tris/HCl/0.1 mM-EDTA/1 mM-cysteine hydrochloride, pH 6.7 (lanes 1-6; '10 mM-Tris') or in the same buffer containing 1 M-NaClO₄ (lanes 7-11; '1 M-NaClO₄'). All other conditions were the same as in Fig. 1. The major resistant peptide p1 and the minor one p2 are indicated.

Fig. 3. Time course of digestion of purified H5 with subtilisin

Purified H5 was incubated with subtilisin at 37 °C in 10 mM-Tris/HCl, pH 8.0, at 5.5 μg/ml (lanes 1-6; '10 mM-Tris') or in 1 M-NaClO₄/10 mM-Tris/HCl, pH 8.0, at 8 μg/ml (lanes 7-11; '1 M-NaClO₄'). In lanes 12-14 ('TFE'), H5 was digested with 350, 700, and 1500 μg of subtilisin/ml in 30% trifluoroethanol/50 mM-Tris/HCl, pH 8.0. All other conditions were the same as in Fig. 1. The single resistant peptide s1 is indicated.

Fig. 4. Time course of digestion of purified H5 with elastase

Purified H5 was incubated with elastase at 37 °C in 10 mM-Tris/HCl, pH 8.8, at 28 μg/ml (lanes 1-6; '10 mM-Tris') or in 1 M-NaClO₄/10 mM-Tris/HCl, pH 8.8, at 240 μg/ml (lanes 7-11; '1 M-NaClO₄'). All other conditions were the same as in Fig. 1. The resistant peptides e1 and e2, visible in lanes 10 and 11, are indicated.

25 °C for trypsin; 20 mM-Tris/HCl (pH 6.7)/150 mM-NaCl/2 mM-MgCl₂/0.1 mM-EDTA/0.5 mM-cysteine hydrochloride, 25 °C for papain; 30 mM-Tris/HCl (pH 8.0)/120 mM-NaCl/3 mM-MgCl₂, 37 °C for subtilisin; 30 mM-Tris/HCl (pH 8.8)/120 mM-NaCl/3 mM-MgCl₂, 37 °C for elastase. Aliquots (225 µl) of nuclear suspension (12.5 A₂₆₀ units/ml) were incubated with 25 µl of 10-times-concentrated proteinase solution. At the times indicated, 25 µl of 55% HClO₄ were added [9]. After 20 min on ice, the precipitated DNA and core histones were removed by centrifugation (12000 g, 10 min). Extracted histones H1a, H1b, and H5 and proteolytic products therefrom were precipitated from the supernatant with trichloroacetic acid [20% (w/v)] on ice for 30 min, pelleted by centrifugation (12000 g, 10 min), washed four times with ice-cold acetone and dried.

Microsequencing, purification of proteolytic fragments and analysis of amino acid composition

After electrophoresis, polypeptides were electroblotted overnight on to Immobilon filters (Millipore) in 10% (v/v) methanol/0.1% SDS/50 mM-boric acid, pH 9, and revealed by staining with Coomassie Blue R-250 (0.1%) for 5 s and destaining in methanol/acetic acid/water (4:1:5, by vol.). The bands corresponding to the indicated fragments were cut out and subjected to *N*-terminal sequence analysis on an Applied Biosystems 470A protein gas-phase sequencer coupled to an on-line amino acid phenylthiohydantoin derivative analyser.

Fragments in larger-scale proteolytic digests were separated by h.p.l.c. on an RP18 column (Synchropak S300; 250 mm × 4.6 mm; from Bischoff) or a C₈ column (wide pore octyl; 250 mm × 4.6 mm; 5 µm particle size; from J. T. Baker) with a guard column (Bischoff), using an LKB Uvicord S h.p.l.c. system.

Fragments were eluted (flow rate 0.75 ml/min) with a gradient of acetonitrile [5–60% (v/v) in 55 min] in aq. 0.1% (v/v) trifluoroacetic acid, and 0.75 ml fractions were collected. Fractions containing limit peptides were freeze-dried, hydrolysed with either 1 µg of leucine aminopeptidase (Sigma; type VI) in 50 µl of 200 mM-sodium phosphate buffer, pH 7.0, for 24 h at 37 °C, or 200 µl of 6 M-HCl for 22 h at 105 °C and analysed for amino acid composition.

RESULTS

The strategy employed in the present experiments was to subject histone H5 to limited proteolysis, then to identify those peptide bonds of the protein that were highly susceptible to hydrolysis and thereby to identify the regions that were stable as the protein was degraded. The enzymes used were trypsin and three endoproteases, which preferentially cleave unstructured protein regions (papain, subtilisin and elastase). Time courses of H5 degradation were performed by incubating purified H5 and the proteinase for various lengths of time. The digestion fragments were then separated by SDS/PAGE. This approach has been used to identify and isolate structured regions (domains) in many proteins (see [10] and [11] for two recent examples). Significantly, we also analysed the structure of H5 in nuclei, presumably its '*in vivo*' structure, by subjecting erythrocyte nuclei to limited proteolysis followed by extraction with HClO₄.

Digestion of purified H5

In a first series of experiments, purified histone H5 and the proteolytic enzyme were incubated in 10 mM-Tris/HCl, conditions under which H5 is unfolded [4]. The time course of

Fig. 5. Time courses of digestion of H1a, H1b, and H5 in erythrocyte nuclei with trypsin, papain, subtilisin and elastase

Isolated erythrocyte nuclei (625 µg of DNA/ml of nuclear suspension) were incubated for the indicated times (min) with tosylphenylalanylchloromethane-treated trypsin (33 µg/ml) at 25 °C in 30 mM-Tris/HCl (pH 8.0)/120 mM-NaCl/3 mM-MgCl₂ (a), with activated papain (200 µg/ml) at 25 °C in 20 mM-Tris/HCl (pH 6.7)/150 mM-NaCl/2 mM-MgCl₂/0.1 mM-EDTA/0.5 mM-cysteine hydrochloride, (b), with subtilisin (120 µg/ml) at 37 °C in 30 mM-Tris/HCl (pH 8.0)/120 mM-NaCl/3 mM-MgCl₂ (c), and with elastase (225 µg/ml) at 37 °C in 30 mM-Tris/HCl (pH 8.8)/120 mM-NaCl/3 mM-MgCl₂ (d). Histones H1a, H1b, and H5 and proteolytic peptides were extracted with 5% HClO₄, precipitated with trichloroacetic acid, and analysed by SDS/15%-PAGE. All other conditions were the same as in Fig. 1. 'Early'-appearing bands that are probably derived from H1a and/or H1b are labelled with dots. The resistant peptides t3, p1, p2, s1, e1, e2 and e3 are indicated. The triangles designate four intermediates appearing during digestion with papain (b) that have electrophoretic mobilities very similar to those of four intermediates appearing during subtilisin digestion (c).

digestion with trypsin shows multiple bands (Fig. 1, lanes 1–6) with no sign of a stable limit digestion product. The position of the previously reported limit trypsin cleavage product (residues 22–100) under structuring conditions is indicated by an arrowhead [3]; see also below and Fig. 1, lanes 10 and 11). A distinct band of any significant intensity and stability representing this or a similar sized cleavage product was not detectable (lane 6). Similar results were obtained by digestion in 10 mM-Tris/HCl with papain, subtilisin and elastase (Figs. 2, 3 and 4, lanes 1–6). A smear of final cleavage products was generated by subtilisin and elastase (Figs. 3 and 4, lanes 6).

Although previous experiments had shown that the central sequence comprising residues 22–100 of H5 is folded in 0.15 M-NaCl and in 0.2 M-K₂SO₄ into a globular domain [3], more recent c.d. measurements indicated that helix formation can be induced in the C-terminal tail of H5 in 1 M-NaClO₄ [6]. This prompted us to repeat the digestion experiments under these ionic-strength conditions. Fig. 1, lanes 7–11, shows a time course of digestion of H5 with trypsin in 1 M-NaClO₄. After generation of a number of intermediates, two limit digestion products, t2 and t3, were produced (lanes 10 and 11). The faster migrating one of these, t3, was somewhat more resistant to proteolytic cleavage than the slower one. This pattern of cleavage products was identical with that reported for digestion with trypsin in 0.2 M-K₂SO₄ [3]. Blotting on to Immobilon filters and microsequencing of the N-terminal nine residues yielded a single major sequence (Ser-Ala-Xaa-Pro-Thr-Tyr-Xaa-Glu) for both peptides. They thus start at residue 22, as deduced from the known amino acid sequence of H5 [12]. As judged by amino acid composition, the C-termini of peptides t2 and t3 are Lys-109 and Lys-100 respectively (or close to these) (Table 1). These results show that peptides t2 and t3 are identical with the major stable products of limited proteolysis with trypsin in 0.2 M-K₂SO₄, designated 'LGH5' and 'GH5' [3].

Digestion of H5 in 1 M-NaClO₄ was then continued with three enzymes that preferentially cleave unstructured regions. Fig. 2, lanes 7–11, shows a time course of digestion with papain. Interestingly the pattern of lower-molecular-mass cleavage fragments bears more resemblance to that generated by trypsin in 1 M-NaClO₄ than to that produced by papain in 10 mM-Tris/HCl. One major stable limit cleavage product, p1, was generated and apparent from 10 to at least 120 min of the time course. A minor limit cleavage product, p2, that has an electrophoretic mobility similar to that of fragment t2, was also produced. The relative (to histone H4) electrophoretic mobility of peptide p1 indicated that it is larger than t3 (compare lanes 11 in Figs. 1 and 2 with the marker histones in lanes H). Microsequencing showed that p1 contains the N-terminal sequence Ala-Ser-Arg-Arg-Ser-Ala-Ser, and thus starts at residue 18. The amino acid composition of p1 (Table 1) shows that it contains 11 arginine and 14 serine residues (observed 10.53 and 13.54 residues respectively). Since the determination of arginine in particular was precise enough to distinguish between 10 and 11 residues, we conclude that papain preferentially cleaved on the C-terminal side of Arg-114 [12]. Thus the main difference between p1 and t3 is an extension of the C-terminus of p1 by 14 amino acids. Digestion of H5 with subtilisin generated the single, strikingly stable, limit cleavage product s1, detectable from 10 to at least 120 min of the time course (Fig. 3, lanes 7–11). The relative (to H4) electrophoretic mobility of s1 indicates that s1 and p1 have similar sizes, both higher than the major limit trypsin product. Peptide s1 contains the N-terminal sequence Arg-Arg-Ser-Ala-Ser-His-Pro-Thr-Tyr, thus starting at residue 20; the C-terminus of s1 is the same as (or close to) that of peptide p1, as deduced from amino acid composition (Table 1) [12]. Finally, H5 was digested with pancreatic elastase, generating the major stable limit cleavage product e1 and the minor one e2 (Fig. 4, lanes 7–11). The

electrophoretic mobility of peptide e1 is similar to that of p1 and s1, and the mobility of peptide e2 similar to that of t2 and p2. Peptides e1 and e2 contain the N-terminal sequence Ser-Ala-Ser-His-Pro-Thr-Tyr, and thus start at Ser-22; the C-terminus of e1 is defined by cleavages at Arg-113 and Arg-114 (Table 1) [12].

In summary, digestion with papain, subtilisin and elastase generates major stable limit cleavage products (p1, s1 and e1) whose electrophoretic mobilities are slightly lower than that of histone H4, whereas that of the limit trypsin cleavage product t3 (named 'GH5' in [3]) is significantly higher. The N-termini of peptides p1, s1 and e1 are located between residues 18 and 22, whereas their C-termini arise from cleavage at, or close to, residue 114. These results therefore define a region between residues 18–22 and residue 114 that is highly resistant to proteolysis by papain, subtilisin and elastase. The minor stable limit digestion products, p2 and e2, as well as the trypsin limit cleavage product, t2, that contain the region between residues 18–22 and approx. 109, indicate a proteolytically susceptible site near residue 109. The stable limit trypsin cleavage product t3 (identical with GH5 [3]), finally, defines a 'core' region, from residues 22 to 100, particularly resistant to proteolysis.

Previous c.d. measurements indicated that trifluoroethanol at concentrations of 50–65 % has an even greater stabilizing effect on the secondary structure of H5 than 1 M-NaClO₄ [6]. However, we found that, in routine time-course experiments, trifluoroethanol [even at concentrations as low as 30 % (w/v)] completely inhibited enzymic activity of all four proteinases within 5 min of their addition (results not shown). Since trypsin and subtilisin were inactivated at a lower rate than the other two enzymes, we monitored the digestion of H5 in 30 % trifluoroethanol with these two proteinases in a concentration-dependent mode. Nevertheless, despite the use of very high proteinase concentrations, the digestion level of 'limiting' resistant peptides was not reached (Figs. 1 and 3, lanes 12–14). We could only demonstrate that proteolysis proceeded via discrete intermediates and that the patterns of these were nearly identical with the corresponding ones obtained by digestion in 1 M-NaClO₄.

Digestion of H5 in nuclei

The method of limited proteolysis has the significant advantage that it is also applicable to analysis of the structure of H5 in nuclei, since the digestion products of H5 (and those of H1a and H1b) can be as selectively extracted with 5 % HClO₄ as the parent protein [9]. Nuclei purified from mature chicken erythrocytes were digested for various times with the proteinases used in the studies described above. H1a, H1b and H5, and proteolytic fragments derived therefrom, were extracted with 5 % HClO₄ and subjected to SDS/PAGE. Two factors were helpful in distinguishing between cleavage products derived from H1a and H1b and those derived from H5. First, the band corresponding to H5 dominates those corresponding to H1a and H1b (see Fig. 5), since H5 is the major lysine-rich histone in mature erythrocytes. The major cleavage bands can thus be assumed to derive from H5. Secondly, upon incubation with (at least) papain, subtilisin and elastase, H1a and H1b were digested at a higher rate than H5. Thus the cleavage products of H1a/H1b appeared earlier in the time courses (Figs. 5b, 5c and 5d, lanes 1 and 2 min).

Treatment of erythrocyte nuclei with trypsin rapidly digested H1a, H1b, and H5 into a pattern of multiple intermediates and then gave rise to a highly stable, resistant peptide (Fig. 5a). The pattern of intermediate fragments below the band labelled by an arrowhead shows very high similarity to that generated by digestion of purified H5 in 1 M-NaClO₄. Pairwise comparison indicates that all bands of the 1 M-NaClO₄ digestion pattern in Fig. 1, lanes 7 and 8, are also present in the nuclear digestion pattern, even at similar relative intensities. Thus the bands below

Table 1. Amino acid compositions of peptides t2, t3, p1, s1, e1 and e2 of histone H5

The observed values ('Obs.') in numbers of residues are expressed assuming Ile = 6. The predicted numbers of residues in the indicated sequences ('Pred.') are based on the sequence in [12].

Peptide... Sequence...	Composition (residues/peptide)											
	t2 22-109		t3 22-100		p1 18-114		s1 20-114		e1 22-114		e2 22-109	
	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.
Asp	1.87	2	1.92	2	1.85	2	1.96	2	1.81	2	1.89	2
Asn	0.98	1	1.09	1	0.96	1	1.04	1	0.99	1	1.03	1
Glu	1.89	2	1.92	2	1.84	2	1.86	2	1.87	2	1.84	2
Gln	3.89	4	4.30	4	3.87	4	4.22	4	4.03	4	4.04	4
Ser	13.15	13	12.10	12	13.54	14	13.04	13	12.84	13	13.19	13
His	2.91	3	2.67	3	2.45	3	2.53	3	2.41	3	2.47	3
Gly	8.23	8	6.93	7	8.19	8	7.92	8	7.87	8	7.79	8
Thr	2.21	2	1.97	2	2.20	2	2.10	2	2.22	2	2.21	2
Arg	6.98	7	6.27	6	10.53	11	10.78	11	8.43	9	6.74	7
Ala	10.88	11	10.13	10	13.21	13	11.84	12	11.74	12	10.69	11
Tyr	2.92	3	2.76	3	2.75	3	2.85	3	2.73	3	2.81	3
Met	0.87	1	0.84	1	0.71	1	0.79	1	0.64	1	0.75	1
Val	2.96	3	2.84	3	3.99	4	3.94	4	3.89	4	2.95	3
Phe	0.84	1	0.76	1	0.74	1	0.81	1	0.74	1	0.74	1
Ile	6.00	6	6.00	6	6.00	6	6.00	6	6.00	6	6.00	6
Leu	6.15	6	5.95	6	6.01	6	6.14	6	6.19	6	5.84	6
Lys	12.59	13	8.73	9	13.84	14	13.75	14	13.73	14	13.47	13

the marked band are probably derived from H5. The limit digestion product (termed t3) has the same electrophoretic mobility as the resistant peptide (t3) found after digestion of purified H5 (compare lanes 5-90 min in Fig. 5a with lanes 10 and 11 in Fig. 1 and marker lanes), contains the *N*-terminal sequence Ser-Ala-Ser-Xaa-Pro-Thr-Tyr and, at the *C*-terminus, Lys-100, as deduced by amino acid composition (results not shown). We thus conclude that both peptides are identical. However, it appears that peptide t3 is more resistant to proteolysis when H5 is digested in nuclei, than when purified H5 is digested in 1 M NaClO₄. This may result from the different buffer conditions used or a stabilization of the structure of H5 on binding to chromatin.

On incubation of erythrocyte nuclei with papain, H1a and H1b were digested at a higher rate than H5 (Fig. 5b, lanes 1 and 2 min). Thus the bands labelled with dots are probably cleavage products of H1a and/or H1b, as they appear early in the time course. Digestion with papain generated a highly stable, resistant peptide, p1, that has the same electrophoretic mobility as the limit peptide produced by digestion of purified H5 (Fig. 5b; see Fig. 2, lanes 8-11); it contains the *N*-terminal sequence Ala-Ser-Arg-Arg-Ser-Ala-Ser and, at the *C*-terminus, Arg-114 (amino acid composition not shown). Both peptides are therefore identical. Peptide p1 was partially cleaved into a peptide having an electrophoretic mobility similar to that of p2 from digestion of purified H5 (see Fig. 2, lanes 8-11). We note again that peptides p1 and p2 were more stable when H5 was digested in nuclei compared with digestion of purified H5 in 1 M-NaClO₄.

On incubation of erythrocyte nuclei with subtilisin, H1a and H1b were again digested at a higher rate than H5 (Fig. 5c, lanes 1 and 2 min). Thus the weak early-appearing bands labelled with dots most probably derive from H1a and/or H1b, whereas the prominent late bands originate from H5. The limit cleavage product, s1, shows an electrophoretic mobility similar to that generated by digestion of purified H5 (see Fig. 3, lanes 8-11), contains the *N*-terminal sequence Arg-Arg-Ser-Ala-Ser-His-Pro-

Thr-Tyr and, at the *C*-terminus, Arg-114 (amino acid composition not shown). This limit peptide is therefore identical with the limit peptide generated by digestion of purified H5.

Finally, upon incubation of erythrocyte nuclei with elastase, H1a and H1b were also digested more rapidly than H5 (Fig. 5d, lanes 1 and 2 min). At least the 'early' band labelled with a dot appears to be a cleavage product of H1a or H1b. Three limit digestion products were generated that differed in their stabilities: a first one (e1) of relatively weak stability, a second prominent and stable one (e2) and a third, less abundant, one (e3), which was produced from peptide e2 by further cleavage. Peptides e1 and e2 have electrophoretic mobilities similar to the limit cleavage products (fragments e1 and e2) generated by digestion of purified H5 (see Fig. 4, lanes 10 and 11). Further, peptides e1 and e2 contain the same major *N*-terminal sequence (Ser-Ala-Ser-His-Pro-Thr-Tyr) and the *C*-terminus of peptide e2 is defined by cleavages at Lys-108 and Lys-109 (Table 1). We therefore conclude that peptides e1 and e2 are identical with fragments e1 and e2, respectively, generated by digestion of purified H5.

In summary, limited proteolysis of H5 in nuclei with papain and subtilisin generated the remarkably resistant peptides p1 and s1, which were identical with the corresponding limit peptides produced by proteolysis of purified H5. This demonstrates that the region between residues 18-20 and residue 114 is as efficiently protected in nuclei as in purified H5. Further, the peptides p2, e2, e3 and t3 indicate the proteolytic sensitivity of the peptide chain of H5 in nuclei at residues Lys-109 and Lys-100. And finally, the significant stability of the limit trypsin peptide t3 shows that the globular domain from residues 22 to 100 is also *in situ* a proteolytically highly resistant region of H5.

We also considered the possibility that, in nuclear suspensions, exchanging H5 molecules rather than chromatin-bound H5 were digested. Further, and more likely, the cleavage products of H5 with trimmed or lacking *N*- and *C*-terminal tails might have dissociated from chromatin, so that the final stages of digestion occurred on free molecules. It is known that the globular domains

that the structured region from residues 101 to 114 is also located close to this lysine residue and that both form a functionally important section of the surface of H5. This section may have a role in determining the path of the linker DNA. Since the lysine residues in the sequence from 101 to 114 are only weakly protected by DNA [21], it may be hypothesized that one or more of the three arginine residues are involved in DNA binding. The sequences of chicken, duck and goose histones H5 are distinguished from those of the closely related histone H1^o and other H1 histones by the highly selective exchange of lysine for arginine residues [14]. Since this exchange has been related to the function of H5 to increase chromatin condensation [22,23], the notion of an involvement of arginine rather than lysine residues in DNA binding gains additional support. Interestingly, within the region from residues 101–120, only one arginine residue, namely Arg-113, is conserved among histones H5 from chicken, duck and goose (boxed by a thick line in Fig. 6), but replaced by a lysine residue in H1^o. This arginine residue is thus a candidate for interaction with linker DNA. It may be further noted that several previous studies concluded that only a small fraction of the lysine and arginine residues is involved in phosphate binding [21,23].

Digestion of purified H5 as well as of H5 *in situ* proceeded through a pattern of discrete intermediates. With the exception of digestion with trypsin, the '*in situ*' patterns are not very similar to those obtained with purified H5. This most probably results from the interaction of the tails of H5 with linker DNA and packaging of H5 into the nucleofilament and higher-order chromatin structures. Consistent with this, previous n.m.r. measurements concluded that histone H5 is largely immobilized in chromatin [24,25]. Striking similarities, however, between the *in situ* pattern after digestion with papain and that after digestion with subtilisin are readily detectable. In detail, comparison of Fig. 5(b), lane 5 min, with Fig. 5(c), lanes 10 to 30 min, shows that a group of four bands (labelled by triangles) is common to the intermediate pattern in both digests. This is striking, since after digestion of purified H5 the patterns of intermediates with both enzymes are quite different (Figs. 2 and 3, lanes 8). Though further analysis is needed to show that corresponding peptides are really the same, the possibility is raised that the similar discrete intermediate patterns were generated by a defined interaction of the structured tails of H5 with linker DNA.

Previously, limited proteolysis has been successfully used to purify limiting resistant peptides for further analysis by physical methods [3,5,13]. In an extension of the fruitful combination of these different approaches, an n.m.r. study of the central folded

domain of H5 (between residues 18–22 and residue 114) may provide additional insights into the structure of the region from residues 101 to 114.

We thank Mr. A. Bauche for help with h.p.l.c., Mr. E. Kotthaus for co-operation in the analysis of amino acid composition at an early stage of these studies and Dr. L. Böhm for reading the manuscript. This work was supported by grants (to W.H.S.) from the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Thoma, F., Koller, T. & Klug, A. (1979) *J. Cell Biol.* **83**, 403–426
2. Neelin, J. M., Callahan, P. X., Lamb, D. C. & Murray, K. (1964) *Can. J. Biochem.* **42**, 1743–1752
3. Aviles, F. J., Chapman, G. E., Kneale, G. G., Crane-Robinson, C. & Bradbury, E. M. (1978) *Eur. J. Biochem.* **88**, 363–371
4. Crane-Robinson, C., Danby, S. E., Bradbury, E. M., Garel, A., Kovacs, A.-M., Champagne, M. & Daune, M. (1976) *Eur. J. Biochem.* **67**, 379–388
5. Clore, G. M., Gronenborn, A. M., Nilges, M., Sukumaran, D. K. & Zarbock, J. (1987) *EMBO J.* **6**, 1833–1842
6. Clark, D. J., Hill, C. S., Martin, S. R. & Thomas, J. O. (1988) *EMBO J.* **7**, 69–75
7. Suzuki, M. (1989) *EMBO J.* **8**, 797–804
8. Clark, D. J. & Thomas, J. O. (1986) *J. Mol. Biol.* **187**, 569–580
9. Johns, E. W. (1964) *Biochem. J.* **92**, 55–59
10. Schröter, H., Mueller, C. G. F., Meese, K. & Nordheim, A. (1990) *EMBO J.* **9**, 1123–1130
11. Hoeck, W. & Groner, B. (1990) *J. Biol. Chem.* **265**, 5403–5408
12. Briand, G., Kmiecik, D., Sautiere, P., Wouters, D., Borie-Loy, O., Biserte, G., Mazen, A. & Champagne, M. (1980) *FEBS Lett.* **112**, 147–151
13. Hill, C. S., Martin, S. R. & Thomas, J. O. (1989) *EMBO J.* **8**, 2591–2599
14. Doenecke, D. & Tönjes, R. (1986) *J. Mol. Biol.* **187**, 461–464
15. Doenecke, D. & Tönjes, R. (1984) *J. Mol. Biol.* **178**, 121–135
16. Yaguchi, M. & Roy, C. & Seligy, V. L. (1979) *Biochem. Biophys. Res. Commun.* **90**, 1400–1406
17. Maeder, D. L. & Böhm, L. (1991) *Biochim. Biophys. Acta* **276**, 233–238
18. Allan, J., Mitchell, T., Harborne, N., Böhm, L. & Crane-Robinson, C. (1986) *J. Mol. Biol.* **187**, 591–601
19. Allan, J., Hartman, P. G., Crane-Robinson, C. & Aviles, F. X. (1980) *Nature (London)* **288**, 675–679
20. Staynov, D. Z. & Crane-Robinson, C. (1988) *EMBO J.* **7**, 3685–3691
21. Thomas, J. O. & Wilson, C. M. (1986) *EMBO J.* **5**, 3531–3537
22. Thomas, J. O. & Rees, C. (1983) *Eur. J. Biochem.* **134**, 109–115
23. Kumar, N. M. & Walker, I. O. (1980) *Nucleic Acids Res.* **8**, 3535–3551
24. Smith, R. M. & Rill, R. L. (1989) *J. Biol. Chem.* **264**, 10574–10581
25. Shimidzu, M., Shindo, H., Matsumoto, U., Mita, K. & Zama, M. (1985) *Eur. J. Biochem.* **148**, 431–436