Amino Acid Sequence Analysis of the Asparagine-288 Region of the Carbohydrate Variants of Human Plasminogen†

James R. Powell and Francis J. Castellino*

ABSTRACT: The amino acid sequences of the two major carbohydrate variants in the region of Asn_{288} have been determined. A peptide isolated from plasminogen variant 1, which contains a complex-type Asn_{288}-linked oligosaccharide, was found to possess the amino acid sequence -Ser_{280}-Ala-Gln-Thr-Pro-His-Thr-His-Asn(CHO)-Arg-Thr_{290}-Pro-Glu-, in agreement with the previously published sequence [Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E., & Magnusson, S. (1978) in *Progress in Chemical Fibrinolysis and Thrombolysis* (Davidson, J. F., Rowan, R. M., Samama, M. M., & Desnoyers, P. C., Eds.) Vol. 3, pp 191–209, Raven Press, New York]. A similar peptide isolated from plasminogen variant 2 did not contain oligosaccharide but possessed an amino acid sequence identical with the corresponding variant 1 peptide. Thus, the basis for the lack of the complex-type oligosaccharide in human plasminogen variant 2 does not reside in substitution of essential amino acid residues in the region of the Asn_{288}-linked glycosylation site.

All mammalian plasminogens, so far examined, can be resolved into two variant forms by affinity chromatography on Sepharose–lysine. In the case of rabbit, sheep, and human plasminogens, three species which have been extensively characterized, the two variant forms differ dramatically in their carbohydrate compositions (Castellino et al., 1973; Hayes et al., 1975; Paoni et al., 1977; Summaria et al., 1976), with affinity chromatography form 1 containing significantly more carbohydrate than affinity chromatography form 2. In the human system, the carbohydrate sequences, and their linkage points to human plasminogen, have been determined (Hayes & Castellino, 1979a–c). It was found that the carbohydrate differences between the two plasminogen variants are accounted for by the presence of a complex-type oligosaccharide present on Asn_{288} of variant 1 plasminogen, which is absent on variant 2 plasminogen.

A possible explanation for the lack of glycosylation of variant 2 human plasminogen is the substitution of an essential amino acid at the potential oligosaccharide attachment site, which, if indeed the case, would also offer a genetic basis for the presence of two major plasminogen variants. We have examined this possibility in the present paper by isolation and amino acid sequence analysis of the region of the two major plasminogen forms possessing the actual or potential site of attachment for the complex-type oligosaccharide.

Materials and Methods

Proteins. Human plasminogen variants 1 and 2 were purified by our modification (Brockway & Castellino, 1972) of the Deutsch & Mertz (1970) affinity chromatography technique. The starting material was Cohn fraction III, generously provided by Cutter Laboratories.

Porcine pancreatic elastase was purchased from the Sigma Chemical Co. and treated with 5000 kallikrein inactivator units (KIU) of Trasylol (FBA Pharmaceuticals) per mg of elastase prior to use.

Staphylococcus aureus protease V8 was obtained from Miles Laboratories.

The kringle 1–3 region (K 1–3) of both variants 1 and 2 of human plasminogen, consisting of amino acid residues Tyr_{79}–Val_{337}, Val_{353}, was isolated by affinity chromatography of Sepharose–lysine (Powell & Castellino, 1980) after generation of this peptide fragment by elastolytic digestion of native plasminogen, essentially as described by Sottrup-Jensen et al. (1978). In each case, the yield of K 1–3 from plasminogen was 75–85%.

Carboxamidomethylation (CAM) of K 1–3. The lyophilized K 1–3 fragments from variant 1 and variant 2 plasminogens were separately dissolved in 50 mL of 0.05 M tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl)[1]/6 M Gdn·HCl, pH 8.6, at room temperature, to a final concentration of 2–3 mg/mL, and flushed with N_2. Reduction with DTT was carried out overnight by addition of a 20-fold molar excess of DTT to K 1–3 disulfides. After this time, 1 mL of iodoacetamide (ca. 10% excess over DTT) was added to the solution and the pH maintained at 8.6 by addition of 4 N NaOH, as required. After 30 min, the peptides were dialyzed against several changes of 0.01 M acetic acid (HOAc) and lyophilized.

Preparation and Isolation of Peptides A-1 and A-2. The reduced and carboxamidomethylated K 1–3 fragments from plasminogen variant 1 (CAM-K $1–3_1$) and variant 2 (CAM-K $1–3_2$) were further digested with *S. aureus* protease for cleavage at Asp and Glu residues (Drapeau et al., 1972; Houward & Drapeau, 1972). The lyophilized peptides were separately dissolved in 50 mL of 0.1 M sodium acetate/0.02% NaN_3, pH 4.0 (1–3 mg/mL), and incubated with *S. aureus* protease, at a ratio of 0.01:1 (w/w) protease:peptide. After 24 h at 37 °C, another addition of the same concentration of protease was accomplished. At a time of 24 h later, the digestions were terminated by addition of 0.2 mL of iPr_2PF (1.0 M in 2-propanol), and 30 min later, the solutions were dialyzed against several changes of 0.01 M HOAc in Spectrapor 6 (exclusion size 2000 daltons) dialysis tubing and lyophilized.

The digests were subsequently redissolved in 2.0 mL of 5% HOAc (v/v)/20% sucrose (w/v) and cleared of insoluble material by low-speed centrifugation. The sample was then applied to a 1.5 cm × 120 cm column of Bio-Gel P-6 (Bio-Rad), equilibrated with 5% HOAc at room temperature. Fractions of 1.0 mL were collected at a flow rate of 2–3 mL/h

† From the Department of Chemistry, University of Notre Dame, Notre Dame, Indiana 46556. *Received August 30, 1982.* This work was supported by Grant HL-13423 from the National Institutes of Health.

[1] Abbreviations: DTT, dithiothreitol; iPr_2PF, diisopropyl fluorophosphate; HPLC, high-performance liquid chromatography; TLC, thin-layer chromatography; DABITC, 4-(*N*,*N*-dimethylamino)azobenzene 4′-isothiocyanate; PITC, phenyl isothiocyanate; Tris, tris(hydroxymethyl)aminomethane; Gdn·HCl, guanidine hydrochloride; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

0006-2960/83/0422-0923$01.50/0

upon elution with the same buffer as used for equilibration.

Preparation and Isolation of Peptides B-1 and B-2. The peptides A-1 and A-2 were further fragmented by specific cleavage at tryptophanyl residues by *o*-iodosobenzoic acid (IBA) (Mahoney & Hermodson, 1974; Mahoney et al., 1981). Peptides A-1 and A-2 were dissolved in 1 mL of 4 M Gdn·HCl/80% HOAc. An IBA stock solution containing 0.1–0.25 mM IBA in 4 M Gdn·HCl/80% HOAc was prepared, and redistilled *p*-cresol was added to a *p*-cresol:IBA ratio of 0.2:1 (mol/mol). This IBA solution was allowed to stand for 4 h at room temperature in the dark prior to use. After this time, a sufficient amount of the IBA solution was added to the peptide solution to produce an IBA:peptide ratio of (100–200):1 (mol/mol). The resulting solution was allowed to incubate for 24 h at room temperature in the dark. At this time, a quantity of 80 mg of DTT was added to quench further reaction. The digest was then dialyzed in Spectrapor 6 (exclusion molecular weight of 1000) dialysis tubing for 24 h at 4 °C against a solution of 50% (v/v) HOAc.

The resulting solution was cleared of insoluble material by low-speed centrifugation and subjected to gel filtration on a 1.5 cm × 120 cm column of Bio-Gel P-6, equilibrated with 5% (v/v) HOAc. Fractions of 1 mL were collected at a flow rate of 3 mL/h. The eluate was monitored by both the absorbance at 280 nm and by a fluorescence method. In the latter method, a quantity of 0.05 mL of the fraction of interest was combined with 1.5 mL of a 0.2 M sodium borate buffer, pH 9.2. While the solution was vigorously vortexed, 0.5 mL of fluorescamine (0.2 g/L in acetone) was added. The relative fluorescence of the solution was then determined at 490 nm, with excitation at 390 nm.

The purity of peptides A-1, A-2, B-1, and B-2 was assessed by paper electrophoresis of Whatman 3MM strips, at pH 1.9, in a solvent consisting of HOAc/HCOOH/H_2O (87:25:888 v/v/v) (Atfield & Morris, 1961) and/or, at pH 7.9, in a solvent consisting of 68 mM *N*-ethylmorpholine in HOAc (Guidotti et al., 1962). In general, 5–10 μg of peptide was subjected to the procedure on a 5 cm × 35 cm strip for 60–80 min at 800 V and 4 °C. Subsequent to electrophoresis, the bands were visualized under ultraviolet light after the paper was sprayed successively with 0.2 M sodium borate, pH 9.2, and a fluorescamine solution (200 mg/L in acetone).

Amino Acid Sequencing. The amino-terminal amino acid sequences of peptides A-1 and A-2 were essentially performed by the semimicro method of Peterson et al. (1972), as modified by Powell & Castellino (1981a). The PTH-amino acids were determined by HPLC (Powell & Castellino, 1981a) and fluorescent TLC (Summers et al., 1973).

The peptides B-1 and B-2 were completely sequenced by utilizing the double-coupling procedure with DABITC and PITC, previously described (Chang & Creaser, 1976; Chang, 1977; Chang et al., 1978), with minor operational modifications. Standard solutions of the 11 DABITC-derived amino acids of importance to this study were prepared and analyzed by HPLC, as described by Chang et al. (1980).

Amino-terminal sequences of peptides B-1 and B-2 were also obtained, enzymatically, with porcine kidney aminopeptidase M (Rohm and Hass). The enzyme was dissolved in a buffer consisting of 25 mM Hepes, pH 7.4, and stored in aliquots at −20 °C. A typical time-course digest at 37 °C included 8 nmol of peptide B-1 or B-2, 70 μL of 25 mM Hepes buffer, pH 7.4, 5 μL of norleucine (2.0 nmol/μL), 5 μL of 10 mM $MgCl_2$ in buffer and 1 μL of aminopeptidase M (ca. 1 μg). Aliquots (20 μL) were removed at various time points and analyzed as described below.

In digestions carried out with bakers' yeast carboxypeptidase Y (Sigma), the enzyme was dissolved in H_2O to a final concentration of 0.7 mg/mL and stored in aliquots at −20 °C. A typical digest at 37 °C included 8 nmol of peptide B-1 or B-2, 65 μL of 25 mM NaOAc buffer, pH 5.5, 5 μL of norleucine and 10 μL of carboxypeptidase Y. Aliquots were removed as described above.

Analysis of Amino Acids Liberated from Enzymatic Digests. Here, a micro method was developed suitable for low concentrations of material. The inhibited dried samples (ca. 1.5–2.0 nmol) were converted to fluorescent derivatives by addition of 20 μL of *o*-phthalaldehyde (OPA) solution (0.8 mg/mL OPA, 2.0 mL/L Brij-35, and 2.0 mL/L 2-mercaptoethanol in 1.0 M potassium borate, pH 10.4, and stored in the dark). After 2 min at room temperature, 20 μL of methanol was added, and 5–15 μL of this solution was immediately assayed by HPLC. The OPA-amino acid derivatives were resolved on a 10 × 250 mm Altec Ultrasphere-ODS reverse-phase column (C_{18}) by a two-step linear gradient program, beginning with 70% 0.01 M sodium phosphate buffer (pH 7.0)/30% methanol and increasing to 50% of each phase for 20 min followed by a second increase to 20% sodium phosphate/80% methanol for 5 min, holding those conditions for an additional 10 min at room temperature and a constant flow rate of 1.5 mL/min. The starting conditions were subsequently reset over a 10-min period. The effluent was monitored by a Varian Fluorichrom fluorescence detector having a broad-band 360-nm excitation filter and an emission filter cutoff below 420 nm and was generally operated at low lamp and gain settings of 50–200 arbitrary attenuator units. A standard mixture of 11 amino acids (2.0 nmol/μL of H_2O) was derivatized as stated for the digest samples for identification and quantitation. Time-course graphs of amino acid release were calculated by comparison with the norleucine internal standard and the standard amino acids.

Amino Acid Analysis. Proteins and peptides were analyzed for amino acid content after total hydrolysis in vacuo in 6 N HCl at 110 °C for 24 h. The hydrolyzed samples were dried under reduced pressure, dissolved in sodium citrate buffer (0.2 N sodium), pH 2.2, and injected onto the long column of a Beckman 117 amino acid analyzer. Samples were quantitated by comparison with a standard of 25 nmol each of 17 amino acids. Under the conditions used, glucosamine, when present, was eluted in a well-defined peak in the second buffer between leucine and tyrosine.

Results

The Asn_{288}-linked glycopeptide region of human plasminogen carbohydrate variant 1 was isolated by successive digestions of plasminogen with porcine pancreatic elastase, which is known (Sottrup-Jensen et al., 1978) to produce in high yield the K 1–3 domain of the protein (residues Tyr_{79}–Val_{337}, Val_{353}), followed by digestion with *S. aureus* protease, to produce peptide A-1, and, lastly, with *o*-iodosobenzoic acid, yielding peptide B-1. The elution behavior of *S. aureus* protease digested CAM-K $1–3_1$, on Bio-Gel P6, is shown in Figure 1A. The peptide fraction labeled A-1 was the only glucosamine-containing peptide, and this material was pooled as indicated. Its purity was established by paper electrophoresis as described under Materials and Methods. The amino-terminal sequence of this peptide was determined by the Edman degradation method with identification of the PTH derivatives by both HPLC and TLC. The amino-terminal sequence was found to be NH_2-Leu-Cys(CAM)-Asp. In the sequence information of human plasminogen published by Sottrup-Jensen et al. (1978), this sequence occurs once in the

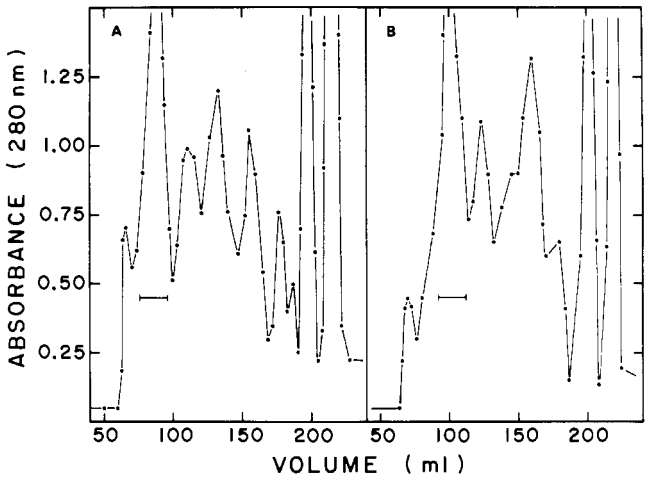

FIGURE 1: (A) Gel filtration elution pattern of *S. aureus* digested carboxamidomethylated kringles 1–3, variant 1, on a 1.5 × 120 cm column of Bio-Gel P-6 (–400 mesh), equilibrated and eluted with 5% acetic acid at room temperature. Fractions of ca. 1.0 mL were collected at a flow rate of about 3 mL/h. (B) Gel filtration of *S. aureus* digested carboxamidomethylated kringles 1–3, variant 2, under the same conditions as (A). The horizontal bars represent the fractions pooled which constitute peptides A-1 and A-2 from human plasminogen variants 1 and 2, respectively.

Tyr_{79}–Val_{337}, Val_{353} region, at residues Leu_{236}-Cys_{237}-Asp_{238}. Thus, the peptide was produced by cleavage at Glu_{235}–Leu_{236}. The cleavage of Asp_{238}–Ile_{239} apparently did not occur. Carboxypeptidase Y digestion indicated that the carboxyl-terminal residue was Glu (0.80 mol/mol). When sequence information of Sottrup-Jensen et al. (1978) is examined, the next available potential Glu cleavage point in sequence is present at position 261. Clearly, this cleavage did not occur, casting some doubt as to the identity of Glu_{261} (this residue is, most likely, Gln, since the total number of Glu residues in the peptide after acid hydrolysis is in agreement with the presumed sequence of the peptide). Proceeding further in the sequence, we determined the next potential cleavage point at Glu_{292}–Asn_{293}. Cleavage here would be consistent with the size and carbohydrate content of peptide A-1. The amino acid composition of peptide A-1, as compared with the composition of a peptide containing residues Leu_{236}–Glu_{292}, is listed in Table I. Clearly, it can be seen that peptide A-1 does consist of this region of plasminogen. The high yield of this peptide (82%) confirms the specificity of the *S. aureus* protease for Glu, under the experimental conditions employed.

Since one Trp residue exists in peptide A-1 (position 279), cleavage with IBA was attempted in order to liberate a small glycopeptide. After reaction with IBA as described under Materials and Methods, the mixture was percolated on a Bio-Gel P-6 column, and the elution pattern is shown in Figure 2A. Two peaks are obtained, peak 1 (B-1) containing glucosamine and peak 2 containing the modified Trp, with absorbance at 280 nm. The yield of peptide B-1, which produced a single band on paper electrophoresis, was 73% from peptide A-1. The amino acid composition of peptide B-1 (after acid hydrolysis) in moles per mole of peptide B-1 was the following: Ala, 1.0; Arg, 1.1; Asx, 0.7; Glx, 1.8; His, 2.1; Pro, 2.2; Ser, 1.1; Thr, 2.9.

Peptide B-1 was subjected to amino acid sequence analysis by the double-coupling method, described under Materials and Methods. Unique identifications were made of the entire 13 amino acid peptide by both HPLC and TLC methodology. The sequence obtained was NH_2-Ser-Ala-Gln-Thr-Pro-His-Thr-His-Asn(CHO)-Arg-Thr-Pro-Glu. The sequence was confirmed by enzymatic digests of peptide B-1. Amino-

Table I: Comparison of the Amino Acid Compositions of Peptides A-1 and A-2 with the Leu_{236}–Glu_{292} Region of Human Plasminogen

amino acid	content (mol/mol) in A-1	A-2	Leu_{236}–Glu_{292}
Ala	2.0	2.0	2
Arg	3.1	3.0	3
Asx	4.1	4.3	4
Cys(CM)	3.8	3.7	4
Glx	4.9	4.7	5
Gly	4.6	4.8	5
His	3.6	3.5	4
Ile	0.9	1.0	1
Leu	2.2	2.0	2
Lys	1.1	0.9	1
Met	0	0	0
Phe	0	0	0
Pro	7.3	7.2	7
Ser	3.5	3.5	4
Thr	8.6	8.7	9
Trp	nd[a]	nd	1
Tyr	2.3	2.1	2
Val	2.9	2.7	3
Glc-NH_2[b]	+	–	4[b]
	82[c]	91[c]	–

[a] nd, not determined. [b] Present in variant 1 only. [c] Percent yield.

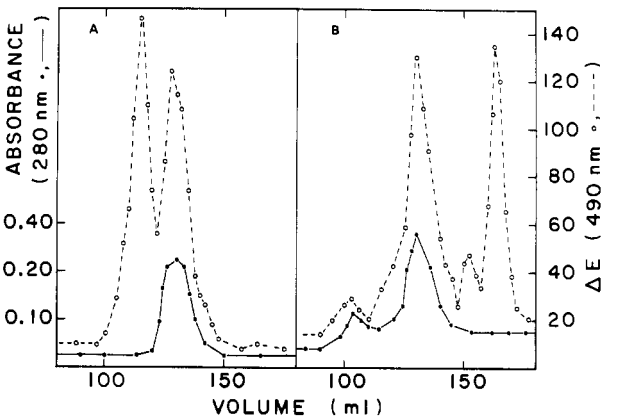

FIGURE 2: (A) Gel filtration profile of IBA-cleaved peptide A-1 from the same column as in Figure 1. (B) Gel filtration profile of IBA-cleaved peptide A-2 under the same conditions as (A). The dashed lines represent the relative fluorescence of individual column fractions assayed with fluorescamine to locate the positions of the peptides B-1 and B-2 from human plasminogen variants 1 and 2, respectively. The solid lines represent the absorbance at 280 nm of the column eluate.

peptidase M digestion of B-1 sequentially liberated Ser (0.98), Ala (0.94), and Gln (0.88) (values are moles per mole of peptide B-1). Trypsin digestion liberated two peptides. One tryptic tripeptide possessed the sequence Thr-Pro-Glu, and carboxypeptidase B treatment of the remaining tryptic peptide resulted in liberation of only Arg (0.92 mol/mol).

Chromatograms of parallel digestions of CAM-K $1–3_2$ with *S. aureus* protease, resulting in peptide A-2, and IBA, resulting in peptide B-2, are shown in Figures 1B and 2B. Despite the lack of carbohydrate on peptide A-2, it nonetheless is the largest peptide present in the *S. aureus* digest of CAM-K $1–3_2$ and is the first major peak of Figure 1B. The yield of A-2 from CAM-K $1–3_2$ was approximately 91%. The amino acid composition of this peptide listed in Table I, its amino-terminal sequence of NH_2-Leu-Cys(CAM)-Asp, and the release of 0.83 mol/mol of Glu by carboxypeptidase Y clearly allow peptide A-2 to be identified as Leu_{236}–Glu_{292}. Interestingly, the amino acid composition of peptide A-2 is identical with that of its counterpart peptide, A-1, suggesting that the amino acid se-

quences of the two peptides may be similar. Cleavage of peptide A-2 with IBA and separation of the peptides by gel filtration, as shown in Figure 2B, allow ready isolation of peptide B-2. In contrast to the behavior of peptide B-1, peptide B-2 elutes as the second peak on this column due to the lack of carbohydrate on the peptide and the corresponding reduction of its molecular weight, when compared to peptide B-1. The yield of peptide B-2 was determined to be approximately 87% from peptide A-2, and a single band was present for B-2 when this peptide was subjected to paper electrophoresis. The amino acid sequence of peptide B-2 was identical with that of peptide B-1 and the amino acid content was found to be the following (values are moles per mole of peptide B-2): Ala, 1.0; Arg, 1.1; Asx, 1.2; Glx, 2.0; His, 1.7; Pro, 2.2; Ser, 1.1; Thr, 2.9. Aminopeptidase M digestion of B-2 sequentially liberated Ser (0.98), Ala (0.90), and Gln (0.87) (values are moles per mole of peptide B-2). Carboxypeptidase B treatment of the tryptic-cleaved peptide B-2 yielded 1.00 mol of Arg, and the remaining tripeptide had the sequence Thr-Pro-Glu. Most interestingly, Asn and Thr were conclusively identified as residues 288 and 290, respectively, in peptide B-2.

Discussion

Human plasminogen is released into plasma as two variant forms, isolatable by affinity chromatography on Sepharose–lysine (Brockway & Castellino, 1972). The first form (variant 1) from the affinity column contains two oligosaccharide chains, a complex glucosamine-containing carbohydrate on Asn_{288} and a trisaccharide unit on Thr_{345}. The second form (variant 2) eluted from the affinity resin contains only the latter oligosaccharide unit (Hayes & Castellino, 1979a–c). The major functional distinction between the two variants involves the strength of their interaction with the ligand ϵ-aminocaproic acid (and its analogues), which is the basis of the chromatographic resolution of the two variants. A previous study from this laboratory (Powell & Castellino, 1981b) has demonstrated that blockage of Asn-linked glycosylation of plasminogen variant 1, in rats and rabbits, by tunicamycin treatment abolishes the functional difference (binding to ϵ-aminocaproic acid) between the two variants. As an important extension of this study, we felt it necessary to determine whether any subtle primary structure differences existed between the two plasminogen variants, which precluded Asn-linked glycosylation of variant 2. Most significantly, we decided to determine whether the necessary Asn-X-Thr/Ser [reviewed by Marshall (1972)] occurred in variant 2. Although the complete amino sequence of human plasminogen has been presented (Sottrup-Jensen et al., 1978), several unresolved problems have led us to reexamine the areas of the sequence of interest to this report. Sottrup-Jensen et al. (1978) obtained sequence data on a combination of variants 1 and 2. In their report, useful data on peptide isolation, peptide yields, and possible microheterogeneity in amino acid residues were not given, severely restricting its usefulness for answering questions of the type posed herein.

When three very selective cleavages in human plasminogen variants 1 and 2 were employed, separately, a 13-residue peptide was isolated, in high yield, from the actual (variant 1) and latent (variant 2) Asn-based oligosaccharide site. The amino acid sequence analysis of each peptide clearly demonstrates that they are identical, and importantly, the variant 2 peptide (B-2) does contain the Asn-X-Thr primary sequence required for glycosylation. Thus, reasons other than critical amino acid substitutions at the oligosaccharide attachment site need to be invoked in order to explain the lack of carbohydrate attachment to Asn_{288} in human plasminogen variant 2.

Both variants of plasminogen are synthesized in the liver parenchymal cells (Bohmfalk & Fuller, 1980), and variant 1 plasminogen is released into plasma at approximately half the rate of variant 2 release (Siefring & Castellino, 1974). Further, the half-life of rabbit and rat plasminogen is of the order of 12–14 h (Siefring & Castellino, 1974). This rapid turnover, coupled with the rate-limiting dolichol phosphate concentration available for protein glycosylation, suggests that plasminogen synthesis and release into plasma may be too rapid for Asn-based glycosylation of all plasminogen molecules. Thus, the presence of two carbohydrate variants is likely due to incomplete Asn-based glycosylation, rather than differences in structure, of the two forms.

Registry No. Plasminogen, 9001-91-6.

References

Atfield, G. N., & Morris, C. (1961) *Biochem. J. 81*, 606–614.

Bohmfalk, J. R., & Fuller, G. M. (1980) *Science (Washington, D.C.) 209*, 408–410.

Brockway, W. J., & Castellino, F. J. (1972) *Arch. Biochem. Biophys. 151*, 194–199.

Castellino, F. J., Siefring, G. E., Sodetz, J. M., & Bretthauer, R. K. (1973) *Biochem. Biophys. Res. Commun. 53*, 845–851.

Chang, J. Y. (1977) *Biochem. J. 163*, 517–520.

Chang, J. Y., & Creaser, E. H. (1976) *Biochem. J. 157*, 77–85.

Chang, J. Y., Brauer, D., & Wittman-Liebold, B. (1978) *FEBS Lett. 93*, 305–314.

Chang, J. Y., Lehmann, A., & Wittman-Liebold, B. (1980) *Anal. Biochem. 102*, 380–383.

Deutsch, D., & Mertz, E. T. (1970) *Science (Washington, D.C.) 170*, 1095–1096.

Drapeau, G. R., Boily, Y., & Houward, J. (1972) *J. Biol. Chem. 247*, 6720–6726.

Guidotti, G., Hill, R. J., & Konigsberg, W. (1962) *J. Biol. Chem. 237*, 2184–2195.

Hayes, M., & Castellino, F. J. (1979a) *J. Biol. Chem. 254*, 8768–8771.

Hayes, M., & Castellino, F. J. (1979b) *J. Biol. Chem. 254*, 8772–8776.

Hayes, M., & Castellino, F. J. (1979c) *J. Biol. Chem. 254*, 8777–8780.

Hayes, M., Bretthauer, R. K., & Castellino, F. J. (1975) *Arch. Biochem. Biophys. 171*, 651–655.

Houward, J., & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3506–3509.

Mahoney, W. C., & Hermodson, M. A. (1974) *Biochemistry 13*, 3810–3814.

Mahoney, W. C., Smith, P. K., & Hermodson, M. A. (1981) *Biochemistry 20*, 443–448.

Marshall, R. D. (1972) *Annu. Rev. Biochem. 41*, 673–702.

Paoni, N. F., Violand, B. N., & Castellino, F. J. (1977) *J. Biol. Chem. 252*, 7725–7732.

Peterson, J. D., Nehrlich, S., Oyer, P. E., & Steiner, D. F. (1972) *J. Biol. Chem. 247*, 4866–4871.

Powell, J. R., & Castellino, F. J. (1980) *J. Biol. Chem. 255*, 5329–5335.

Powell, J. R., & Castellino, F. J. (1981a) *Biochem. Biophys. Res. Commun. 102*, 46–52.

Powell, J. R., & Castellino, F. J. (1981b) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6836–6839.

Siefring, G. E., & Castellino, F. J. (1974) *J. Biol. Chem. 249*, 1434–1438.

Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E.,

& Magnusson, S. (1978) in *Progress in Chemical Fibrinolysis and Thrombolysis* (Davidson, J. F., Rowan, R. M., Samama, M. M., & Desnoyers, P. C., Eds.) Vol. 3, pp 191–209, Raven Press, New York.

Summaria, L., Spitz, F., Arzadon, L., Boreisha, I. G., & Robbins, K. C. (1976) *J. Biol. Chem. 251*, 3693–3699.

Summers, M. R., Smythers, G. W., & Oroszlan, S. (1973) *Anal. Biochem. 53*, 624–628.

Chromatin Structure Differs between Coding and Upstream Flanking Sequences of the Yeast 35S Ribosomal Genes†

D. Lohr

ABSTRACT: Staphylococcal nuclease (EC 3.1.4.7) and DNase I (EC 3.1.4.5) digestion analysis of the nuclear chromatin structure of the yeast 35S rDNA gene shows the presence of typical and homogeneous nucleosome patterns across the coding sequence. These nucleosomal patterns change abruptly, around the site of transcription initiation and upstream in the 5′-flanking sequences, to a unique pattern with both nucleosomal and nonnucleosomal character. The mix arises, at least partly, from heterogeneity within the population of upstream regions; some regions are nucleosomal, but the majority are nonnucleosomal. The nonnucleosomal set of upstream regions appears to be nucleoprotein associated and, in fact, may be an altered nucleosome structure rather than totally restructured. The abruptness of the transition from nucleosome to other structure suggests restricted nucleosome locations in the region around the transcription initiation site of this gene.

Transcription initiation is one facet of the control of eukaryotic gene expression. The specificity for the location and the frequency of transcription initiation resides in the upstream 5′ sequences for genes transcribed by RNA polymerase II (Breathnach & Chambon, 1981) and probably for RNA polymerase I (Grummt, 1981). There are indications that, in addition to the functional distinction between upstream (control) and coding sequences, there are chromatin structural differences between these regions. For example, DNase I and staphylococcal nuclease hypersensitive (cleaved very quickly) sites are found mainly in the upstream rather than coding sequences (cf. Elgin, 1981). Restriction endonucleases can also cleave preferentially in the control region, demarcating an accessible domain of ~400 base pairs (bp)[1] in SV40 which correlates with a nonbeaded region of the minichromosome (Varshavsky et al., 1979) and allowing complete excision of an ~115-bp fragment from the upstream region of a β-globin gene in chicks (McGhee et al., 1981). Distinctions between coding and noncoding regions may even be present in the DNA itself. Keene & Elgin (1981) have shown, by staphylococcal nuclease digestion of protein-free DNA, that noncoding sequences are cleaved preferentially at ~200-bp intervals on average, while adjacent coding regions are not so cleaved.

The SV40 data suggest another possible structural distinction between coding and control sequences, the presence or absence of nucleosomes. In SV40 the "open" region contains control elements (Varshavsky et al., 1979), while the nucleosome domain begins in the coding region. Other studies have detected nucleosomes preferentially located on the noncoding sequences adjacent to genes, cf. *Drosophila* histone genes (Samal et al., 1981) and *Drosophila* heat-shock genes (Wu, 1980). The interest in structural distinctions between control and coding sequences derives from the possibility that such structural distinctions could form some of the basis for the functional distinction, the specificity of transcription initiation.

To study the chromatin structure of individual genes, I developed a technique (Lohr, 1981) which permits analysis of the products of nuclease digestion of the chromatin on a single DNA sequence to the same level of detail and using the same spectrum of approaches as the original work which led to the elucidation of the bulk nucleosomal structure of chromatin. The technique utilizes electrophoretic transfer (Stellwag & Dahlberg, 1980) of DNA patterns to DBM paper (Alwine et al., 1980), which binds the DNA covalently. Thus one can transfer from high-resolution polyacrylamide gels, allowing the detailed analysis of staphylococcal nuclease and DNase I digestion products as double-strand patterns or as single-strand patterns on denaturing gels and permitting analysis of quite small sizes of DNA (down to 80 nucleotides). The extra effort involved in using this technique is justified because the method can produce "structural" information, in addition to "kinetic" information (hypersensitive sites etc.), which is useful but can tell little about why a sequence is preferentially cleaved and thus is difficult to interpret structurally. Application of this method to the region around the 35S rDNA transcription initiation site shows some unique chromatin structural features.

Materials and Methods

Cells were grown to early log (5×10^7/mL) or stationary [$(20–30) \times 10^7$/mL] in YEPD (1% yeast extract, 2% Bactopeptone, and 2% dextrose). Nuclei were isolated, digestions were performed with staphylococcal nuclease or DNase I, and DNA was extracted as described (Lohr et al., 1977b). Electrophoresis on composite nondenaturing or denaturing gels and DNA transfer to DBM paper was exactly as described in Lohr (1981) except the nondenaturing gels were treated in 0.25 M HCl for 15 min to depurinate the DNA as described in Alwine et al. (1980). All gels contained unlabeled PM2–*Hae*III restriction fragments which were visualized by including end-labeled PM2 fragments in the hybridization mix. Hybridizations were done according to Alwine et al. (1980) as described in Lohr (1981).

† From the Chemistry Department, Arizona State University, Tempe, Arizona 85287. *Received August 17, 1982.* This research was supported by Grant PHS-GM27623.

[1] Abbreviations: bp, base pairs; b, nucleotide bases; DBM, diazobenzyloxymethyl.

0006-2960/83/0422-0927$01.50/0