The EMBO Journal vol.7 no.9 pp.2681 – 2686, 1988

Isolation and structural analysis of a peptide containing the novel tyrosyl – glucose linkage in glycogenin

Carl Smythe, F.Barry Caudwell, Michael Ferguson[1] and Philip Cohen

Department of Biochemistry, Medical Sciences Institute, University of Dundee, Dundee DD1 4HN, and [1]Department of Biochemistry, University of Oxford, Oxford, UK

Communicated by P.Cohen

The glucosylation site on glycogenin, the protein primer required for *de novo* glycogen synthesis, has been identified. The glucose is attached at position C1 in a glycosidic linkage with a unique tyrosine, and the sequence surrounding this residue was found to be: His-Leu-Pro-Phe-Ile-Tyr-Asn-Leu-Ser-Ser-Ile-Ser-Ile-Tyr(Glc)-Ser-Tyr-Leu-Pro-Ala-Phe-Lys. The same tyrosine residue is glucosylated whether glycogenin is isolated as a complex with the catalytic subunit of glycogen synthase, or covalently attached to glycogen. The possibility that insulin and growth factors may enhance glycogen synthesis via stimulation of the priming reaction is discussed.

Key words: glucosylation/glycogen initiation/glycogen synthesis/amino acid sequence/mass spectrometry

Introduction

Since the discovery and characterization of the enzymes involved in glycogen biosynthesis, there has been a problem concerning the mechanism by which glycogen molecules originate. The enzyme glycogen synthase is completely inactive unless maltose or a longer oligosaccharide is present, yet enzyme systems for producing such primers have not been identified.

The *de novo* synthesis of protein-bound oligosaccharides from UDP[^{14}C]glucose by enzyme(s) present in a 150 000 *g* pellet from rat liver was first described by Krisman and Barengo (1975) who proposed that glycogen biogenesis was brought about on a protein primer, which acted as a recipient for glucose residues transferred by a putative 'glycogen initiator synthase' to numerous sites on the protein molecule. A prediction of this hypothesis was that glycogen particles should contain covalently bound protein, and such a protein was identified by Kennedy *et al.* (1985), who termed it 'glycogenin'.

Recently, Pitcher *et al.* (1987, 1988) showed that glycogen synthase from rabbit skeletal muscle contained stoichiometric amounts of glycogenin complexed to the catalytic subunit of glycogen synthase and established that two distinct enzymatic activities were required to form the primer for glycogen synthesis. The first step involved the covalent attachment of a single sugar residue to the glycogenin polypeptide catalysed by the 'glycogen initiator synthase' as proposed by Krisman and Barengo (1975), while the second was an autoglucosylation reaction catalysed by glycogenin itself. Glycogenin was shown to be a glucosyl transferase which in the presence of Mn^{2+} or Mg^{2+} and UDP-glucose (glc) autoglucosylated to a stoichiometry of up to ~6 glucose residues per mol to generate the primer that could be elongated by glycogen synthase (Pitcher *et al.*, 1988).

In this paper we establish that glucose is attached covalently to a unique tyrosine residue on glycogenin, a novel glycosidic linkage.

Results

Isolation of [^{14}C]glucopeptides from glycogenin

The glycogen synthase – glycogenin complex (Pitcher *et al.*, 1987) was treated for 30 min with 2 M LiBr, which inactivates glycogen synthase without affecting the glucosyl transferase activity of glycogenin (Pitcher *et al.*, 1988). After removal of LiBr, the glycogenin was allowed to autoglucosylate in the presence of 5 mM Mn^{2+} and 20 μM UDP[^{14}C]glc to an average stoichiometry of ~1 mol glucose/mol protein, digested with elastase and subjected to gel filtration on Bio-Gel P6 (Figure 1) followed by reverse phase HPLC at pH 6.5 using a C_{18} column (Figure 2A).

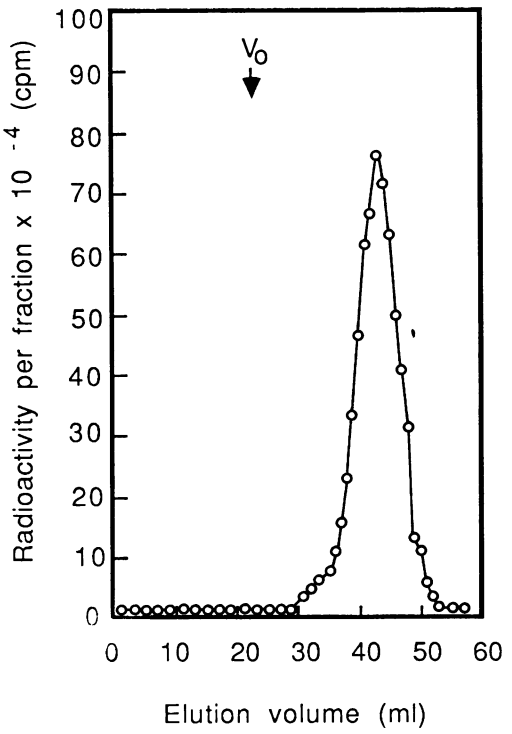

Fig. 1. Gel filtration of [^{14}C]glucosylated peptides on Bio-Gel P6. An elastase digest of the glycogen synthase – glycogenin complex, the latter glucosylated using UDP[^{14}C]glc, was chromatographed on Bio-Gel P6 (43 × 1.5 cm) equilibrated in 0.05 M NH_4HCO_3. The radioactive peak eluted at a V_e/V_o of 2.0 and recovery of radioactivity was 76%. The flow rate was 60 ml/h and fractions of 1 ml were collected.

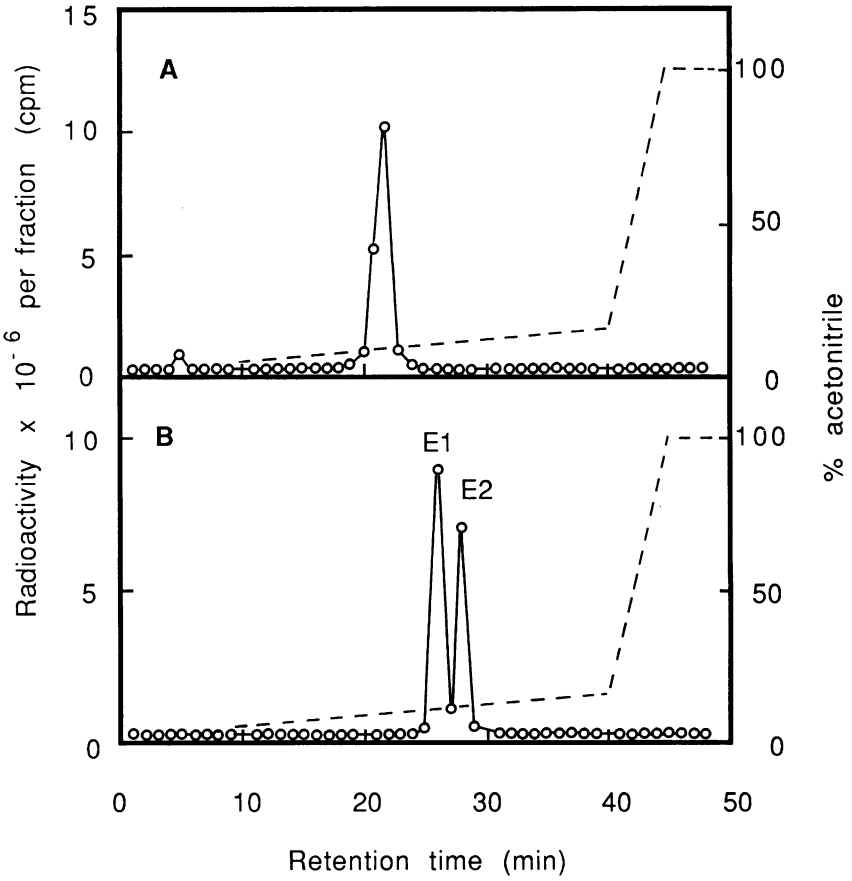

Fig. 2. Isolation of peptides E1 and E2 by reverse phase HPLC (**A**) The radioactive peak from Figure 1 was chromatographed by reverse phase HPLC at pH 6.5 as in Materials and methods. The single radioactive peak eluted at 7.4% (v/v) acetonitrile and recovery of radioactivity was 77%. (**B**) The radioactive fractions from (A) were pooled and rechromatographed at pH 1.9 as in Materials and methods. Peptides E1 and E2 eluted at 9.1% (v/v) and 9.8% (v/v) acetonitrile respectively. Recovery of radioactivity was 72%. The acetonitrile gradient is shown by the broken line.

The [^{14}C]peptide(s) which eluted as a single peak from both columns was rechromatographed on the C_{18} column at pH 1.9, which resolved it into two peaks, E1 and E2 (Figure 2B). Peptide E1 was of a high degree of purity and had the composition Ser_2, Tyr, Ile (Table I). Peptide E2 was less pure, although subsequent sequence and mass spectrometric analyses described below indicated that it had the composition Ser, Tyr, Ile (Table I).

Sequence analysis showed that the primary structure of E1 was Ser-Ile-Xaa-Ser, while E2 was Ser-Ile-Xaa contaminated with the peptide Lys-Met-Leu-Asp-Lys-Glu. In the case of E1 and E2, no Pth-amino acid was observed at the third cycle of Edman reaction and all of the applied radioactivity remained bound to the glass fibre disc, indicating that the [^{14}C]glucosylated amino acid (presumably tyrosine) was not extracted by the sequencer solvents.

The glycosylated amino acid in glycogenin is tyrosine

Substitution reactions involving the phenolic side chain of tyrosine require the presence of a free hydroxyl group (Mayberry *et al.*, 1965) and we utilised this fact to establish that glycosylation occurs via a tyrosine in peptides E1 and E2. Neither peptide E1 nor E2 could be iodinated using the chloramine T method, whereas a synthetic peptide corresponding to unglucosylated E2 was iodinated to a stoichiometry of 1 mol/mol under identical conditions (Table II). Furthermore, when peptide E1 was treated for 2 h at 100°C with 1 N HCl to remove the carbohydrate side chain, it could be iodinated to a stoichiometry of ~1 mol/mol. These results established that the sugar residue was linked covalently to the tyrosine residue in E1 and E2.

Isolation of a [^{14}C]glycosylated tryptic peptide

It was necessary to determine whether glycogenin contained two glycosylation sites with very similar peptide sequences, or if the two peptides E1 and E2 were generated by differential cleavage at a single site. Glycogenin, autoglucosylated as before, was digested with trypsin and peptides separated by reverse phase HPLC at pH 1.9. A single major peak of ^{14}C radioactivity was obtained (Figure 3) which accounted for 79% of the radioactivity recovered from the column. Sequence analysis showed that the primary structure of the [^{14}C]glycosylated peptide was:
His-Leu-Pro-Phe-Ile-Tyr-Asn-Leu-Ser-Ser-Ile-Ser-Ile-Xaa-Ser-Tyr-Leu-Pro-Ala-Phe-Lys.
No Pth-amino acid was identified at cycle 14 suggesting that this was the location of the glycosylated tyrosine. Residues 12–15 and 12–14 corresponded to peptides E1 and E2 respectively.

Table I. Amino acid compositions of peptides E1 and E2 containing the glucosylation site of glycogenin

Amino acid	E1*	E1**	E2*,[a]	E2**
Asx	0.26		0.79	
Glx	0.26		0.67	0.27
Ser	1.80(2)	1.95(2)	1.26(1)	1.06(1)
Gly	0.44		0.59	
His				
Arg			0.29	
Thr			0.26	
Ala				
Pro		0.22	0.26	1.16
Tyr	0.98(1)	0.75(1)	1.12(1)	0.7(1)
Val				
Met		0.25	0.73	
Ile	1.00(1)	1.00(1)	1.00(1)	1.00(1)
Leu	0.34			0.99
Phe				
Lys	0.36		1.2	
Total	4	4	3	3
[^{14}C]glucose	1.3(1)		1.1(1)	

Peptides were hydrolysed for 16 h at 110°C in 6 N HCl–2 mM phenol. Values for serine and threonine were corrected for 10 and 5% destruction respectively, and impurities below 0.2 mol are omitted Numbers in parentheses indicate residues determined by sequencing (E1*, E2* and E1**) or by FABMS (E2**).
*[^{14}C] labelled peptides isolated from the glycogen synthase–glycogen complex.
**unlabelled peptides isolated from glycogenin glucosylated *in vivo*.
[a]The preparation also contained the peptide Lys-Met-Leu-Asp-Lys-Glu.

Table II. Iodination of peptides derived from glycogenin

Peptide	Stoichiometry of iodination
	(mol I/mol peptide)
E1	<0.01
E2	<0.01
deglycosylated	
E1	1.2
synthetic	
E2	1.0

Analysis by fast-atom bombardment mass spectrometry of glycopeptides from glycogenin glycosylated in vivo

Glycogen was purified as in Pitcher *et al.*, (1987) and digested exhaustively with α-amylase and glucoamylase to obtain glycogenin. After further digestion with elastase, peptides were subjected to reverse phase HPLC at pH 6.5 followed by rechromatography at pH 1.9. A single peak of carbohydrate was obtained from the column run at pH 6.5 (Figure 4A) which eluted at the same position as the peak of ^{14}C radioactivity obtained from the glycogenin complexed to glycogen synthase after labelling with UDP[^{14}C]glc *in vitro* (Figure 2). After rechromatography at pH 1.9 (Figure 4B), peptide E1 was identified by amino acid analysis (Table I) and E2 by fast-atom bombardment mass spectrometry (FABMS) (Figure 5).

Two major signals with molecular masses (MH^+) of 634 and 796 were obtained when E1 was subjected to FABMS

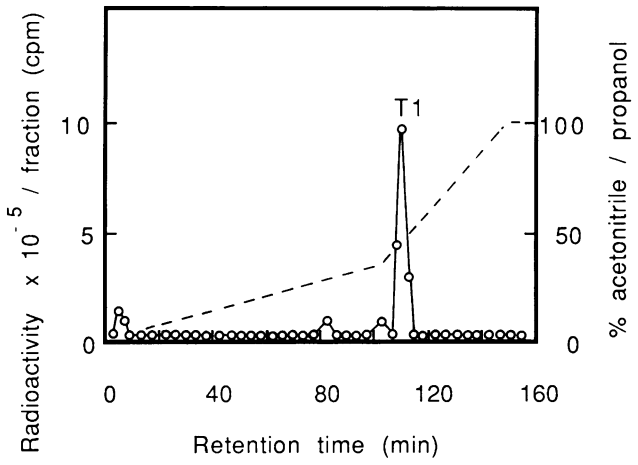

Fig. 3. Isolation of tryptic peptide T1 by reverse phase HPLC at pH 1.9. A tryptic digest of the glycogen synthase–glycogenin complex, the latter glucosylated using UDP[^{14}C]glc was chromatographed as described in Materials and methods. The single major radioactive peak eluted at 30% acetonitrile:10% propanol. Recovery of radioactivity was 40%. The acetonitrile/propanol gradient is shown by the broken line.

(Figure 5A). These values, which are 162 and 324 mass units greater than that expected for the peptide Ser-Ile-Tyr-Ser, indicated that E1 was a mixture of two peptides containing one and two unsubstituted hexose residues. The two minor signals at 654 and 816 mass units are likely to be the sodium adducts of the monoglucosylated and diglucosylated peptides.

Peptide E2 gave a single major signal MH^+ at 543, with a minor signal at 565 corresponding to the sodium adduct (Figure 5B). This is the value expected for the monoglucosylated form of the peptide Ser-Ile-Tyr.

Identification of the sugar residues in peptides E1 and E2 and nature of the carbohydrate–peptide linkage

The glycopeptides E1 and E2 were analysed for monosaccharide content by gas chromatography–mass spectrometry following methanolysis and trimethylsilylation as described by Ferguson *et al.* (1988). These analyses indicated that glucose was the only monosaccharide present, and together with the FABMS data establish that the hexose linked covalently to peptides E1 and E2 is glucose.

The glucopeptides were also subjected to methylation analysis as described by Ferguson *et al.* (1988). In both cases, the only partially methylated alditol acetate derivatives observed were 1,5-diacetyl,2,3,4,6-tetramethyl glucitol and 1,4,5-triacetyl,2,3,6-trimethyl glucitol in a molar ratio of 1:1.3 for E2, and 1:2.4 for E1, arising from the non-reducing terminal Glc and 4-*O*-substituted Glc respectively. The presence of these, and no other, derivatives demonstrates that glucose must be attached covalently to the tyrosine via the 4- or 1-position. The former would require a structure of the type Glc1-*O*-Glc4-*O*-Tyr which would be consistent with the observed methylation products. However, this structure can be eliminated since digestion with amylase and glucoamylase could not have generated the monoglucosylated peptide observed in FABMS (Figure 5). The most likely linkage between glucose and the peptide is a Glc1-*O*-Tyr glycosidic linkage with subsequent glucose residues being linked (Glc1-*O*-4Glc)$_n$.

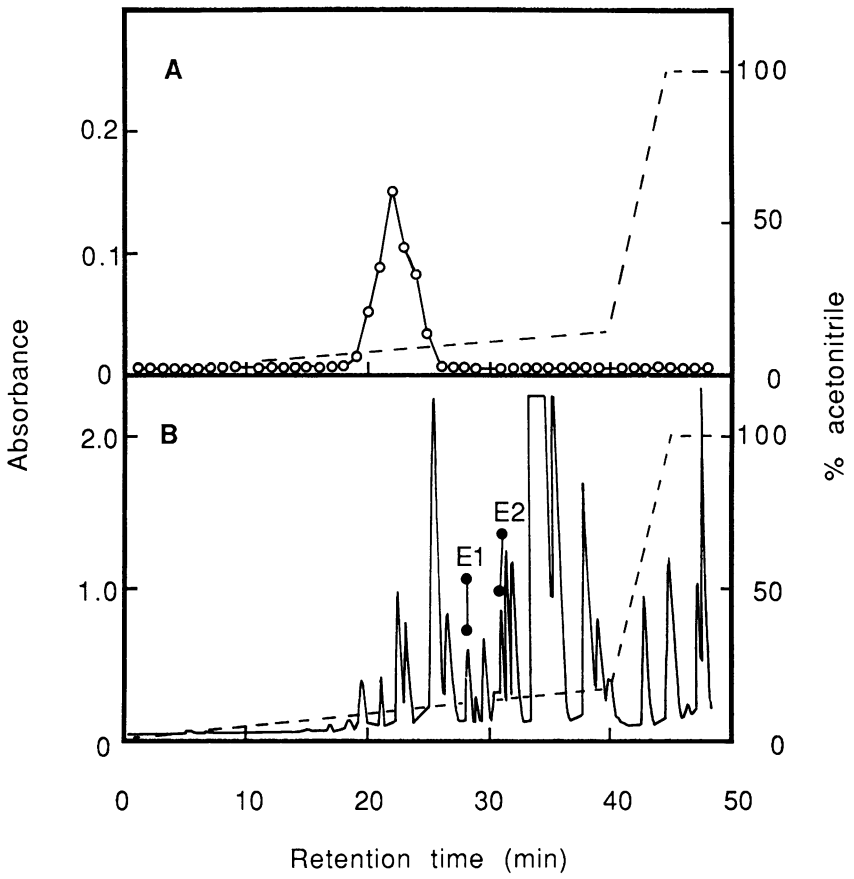

Fig. 4. Isolation of peptides E1 and E2 from glycogenin glucosylated *in vivo*. Reverse phase HPLC at pH 6.5 (**A**) and pH 1.9 (**B**) was performed sequentially Figure 2. In A, aliquots (20%) of each fraction (1 ml) were dried, resuspended in water and assayed for carbohydrate (0–0). The carbohydrate containing fr from (A) were pooled and peptides E1 and E2 resolved by rechromatography at pH 1.9 (B).

Although the FABMS data (Figure 5) clearly showed that the major molecular species in E1 and E2 were Ser-Ile-Tyr($Glc_{1 \text{ and } 2}$)Ser and Ser-Ile-Tyr(Glc_1) respectively, the methylation analyses revealed that peptides bearing larger numbers of glucose units must also be present to account for the stoichiometry of the 4-*O*-substituted glucose derivatives. Traces of highly glucosylated peptides were presumably not observed in FABMS, either because of their low abundance, or due to suppression (Poulter *et al.*, 1987) by the monoglucosylated and diglucosylated derivatives.

Discussion

In this paper, we describe the isolation and structure of a novel glycosylation site on glycogenin, an intracellular protein required for the initiation of glycogen biogenesis. Very few intracellular glucosylation reaction have been characterized, and to our knowledge, the only other example is the *O*-linked *N*-acetylglucosamine residues that are attached to certain proteins on the cytoplasmic face of the nuclear membrane (Holt and Hart, 1986). The results presented here establish that glycogenin is glucosylated at a single site *in vivo* and that the first sugar residue attached to the protein is an unsubstituted glucose molecular that is attached via a glycosidic linkage at the C1 position. The latter observation was anticipated, since it is well known that glycogen molecules do not contain a free reducing end (Rodriguez and Fliesler, 1988). This carbohydrate–protein linkage constitutes the single initiation site for the biogenesis of glycogen. It is to this site that glucosyl residues are subsequently added in an $\alpha(1-4)$ linkage, initially by glycogenin itself (Pitcher *et al.*, 1988) and thereafter by the action of glycogen synthase, with $\alpha(1-6)$ branch points being introduced by branching enzyme.

The linkage of glucose to a tyrosine residue would appear to be unique and confirms an observation by Rodriguez and Whelan (1985) who reported that two tyrosine residues in glycogenin could not be iodinated or nitrated. We, however, have been unable to find any evidence for a second glucosylation site, unless the long tryptic peptide sequence surrounding the tyrosine is repeated in the molecule. The presence of a single glucosylation site is consistent with the finding that each muscle glycogen β-particle contains one molecule of glycogenin (Kennedy *et al.*, 1984).

Glycogen joins a growing number of macromolecules whose synthesis is initiated on a protein primer. These include certain viral RNA and DNA genomes (Wimmer, 1982) as well as long chain fatty acids (McCarthy and Hardie, 1984). The enzyme that catalyses the formation of a tyrosyl–glucose bond ('glycogen initiator synthase') has still to be identified. Its detection may require the isolation of glycogenin molecules that are not glucosylated, but it will be of interest to examine whether a synthetic peptide corresponding to the glucosylation site sequence can be used to detect the putative 'glycogen initiator synthase'.

The potential importance of the glycogenin molecule in

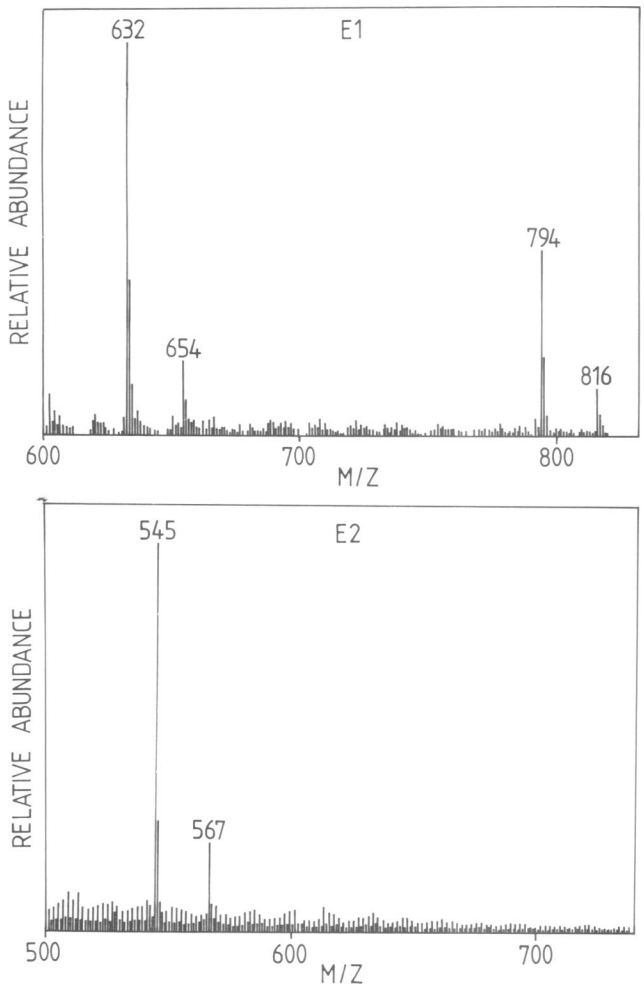

Fig. 5. FABMS analysis of peptides E1 and E2. The relevant regions of the mass spectra are shown. Further details are given in Materials and methods and the spectra are interpreted in Results.

the regulation of glycogen metabolism is now beginning to emerge. The maximum number of glycogen molecules in the cell is determined by the number of glycogenin molecules, and probably more importantly, by the number of glycogenin molecules sufficiently glucosylated to act as a substrate for glycogen synthase. It is intriguing that glycogenin is glucosylated on a tyrosine residue, and that insulin and other growth factors which stimulate glycogen synthesis (Chan and Krebs, 1985) do so by activating protein tyrosine kinases associated with their receptors (Hunter, 1985). Is it possible that phosphorylation of the tyrosine residue on glycogenin is a prerequisite for its glucosylation? Even if this were not the case, these hormones may stimulate glycogen synthesis at the level of initiation, as well as at the elongation step catalysed by glycogen synthase. In this way, they would promote the formation of new glycogen molecules, as well as the enlargement of pre-existing ones.

Materials and methods

Materials

UDP-glc and chloramine T were obtained from Sigma Chemical Co. (Poole, UK), UDP[^{14}C]glc and [^{125}I]NaI from Amersham International (Amersham, UK), trypsin (treated with tosylphenylchloromethylketone) from Worthington Diagnostic Systems (New Jersey, USA), elastase and glucoamylase from Boehringer (Mannheim, FRG) and Bio-Gel P6 from BioRad Laboratories (Watford, UK). Glycogen synthase was purified from rabbit skeletal muscle (Nimmo *et al.*, 1976) and glycogenin was isolated from purified glycogen (Pitcher *et al.*, 1987). Human salivary α-amylase was purified (Embi *et al.*, 1979). The peptide Ser-Ile-Tyr was synthesized on a Cambridge Biosciences peptide synthesizer.

Glucosylation of glycogenin

The glycogen synthase–glycogenin complex (10 mg/ml) was treated for 30 min at 0°C with 2 M LiBr to inactivate the glycogen synthase catalytic subunit, and subjected to gel-filtration on a column of Sephadex G50 Superfine (20 × 1 cm) equilibrated in 100 mM MES, pH 7.0. This preparation was incubated for 30 min at 30°C at a final concentration of 0.38 mg/ml in 4 ml of 100 mM MES pH 7.0, 5 mM dithiothreitol, 5 mM $MnCl_2$, 10 μM UDP[^{14}C]glc (330 mCi/mmole) and the reaction terminated by addition of trichloroacetic acid (TCA) to 2% (w/v).

Digestion with proteinases

The TCA precipitate obtained by centrifugation of the above suspension for 1 min at 4000 *g* was washed twice with water and suspended by sonication in 1 ml of 50 mM NH_4HCO_3. Proteolytic digestion with either elastase or trypsin was carried out for 16 h at 37°C at a weight ratio proteinase:protein of 1:20. Glycogenin isolated from purified glycogen was digested in an identical manner.

Isolation of [^{14}C]glucosylated peptides generated by elastase digestion

Peptides were initially subjected to gel-filtration on Bio-Gel as described in the legend to Figure 1. The [^{14}C]fractions (Figure 1) were taken almost to dryness in a vacuum concentrator (Savant Instruments), pooled and the pH adjusted to 6.5 with 1.0 M acetic acid. Peptides were then resolved by HPLC on a Vydac 218TP54C_{18} column (Phase Separations Group, Hesperia, CA, USA) equilibrated in 10 mM ammonium acetate pH 6.5, using a linear gradient from this solution to acetonitrile. The ^{14}C peak (Figure 2A) was pooled and taken almost to dryness. It was then diluted into 0.1% (v/v) trifluoroacetic acid (TFA) pH 1.9 and rechromatographed on the Vydac C_{18} column equilibrated in 0.1% TFA, using a linear acetonitrile gradient.

Isolation of a [^{14}C]glucosylated tryptic peptide

The proteolytic digest was dried as before, redissolved in 0.1% (v/v) TFA and chromatographed on the Vydac C_{18} column. Peptides were eluted with a gradient of acetonitrile:propanol [3:1 (v/v)].

Amino acid and sequence analysis

Amino acid compositions were determined using a Waters 'PICOTAG' System as described in Holmes *et al.* (1986). Primary structures of peptides were determined using an Applied Biosystems 470A gas-phase sequencer, equipped with an on-line reverse phase chromatography system for identification of Pth-amino acids.

Fast atom bombardment mass spectrometry

FAB mass spectra were recorded in the positive ion mode at an 8 kV accelerating voltage, using a VG70-250SE mass spectrometer (VG Analytical, Manchester, UK) equipped with an integrated data acquisition system. An Ion-Tech source operating at 8 kV was employed to generate the fast Xenon beam. Peptide samples (1–2 nmol) in 10% acetic acid were concentrated to a small volume on the stainless steel probe tip and 0.5 μl of matrix (dithiothreitol:dithioerythritol (3:1) in 1% HCl) was added before insertion into the machine.

Iodination of peptides

Tyrosine-containing peptides were dissolved in 0.3 ml of 0.35 M sodium phosphate pH 7.5. [^{125}I]NaI (2 nmol, 100 mCi/mmol) was added, followed by 0.1 ml of chloramine T (2 mg/ml) in 0.05 M sodium phosphate pH 7.5. After 2 min at ambient temperature, the reactions were stopped by addition of 0.4 ml of sodium bisulphite (2.4 mg/ml) in 0.05 M sodium phosphate pH 7.5. An additional 0.6 ml of water was added and following application to a Sep-Pack C_{18} cartridge (Waters Associates), the column was washed with 20 ml of 0.1% TFA followed by 20 ml of 2% (v/v) acetonitrile in 0.1% TFA. Radiolabelled peptides were eluted with 40% acetonitrile in 0.1% TFA, dried, repurified by reverse phase HPLC at pH 1.9 and quantitated by amino acid analysis.

Carbohydrate assay

This was performed according to the method of Dubois *et al.* (1956).

Acknowledgements

We wish to thank Dr D.Campbell for assistance with peptide sequencing and Dr M.Hubbard for critical reading of the manuscript. This work was supported by the British Diabetic Association and by a Programme Grant and Group Support from the Medical Research Council, London. The fast atom bombardment mass spectrometer was provided by a grant from the Wellcome Trust.

References

Chan,C.P. and Krebs,E.G. (1985) *Proc. Natl. Acad. Sci. USA.*, **82**, 4563–4567.

Dubois,M., Gilles,K.A., Hamilton,J.K., Rebers,P.A. and Smith,F. (1956) *Anal Chem.*, **28**, 350–356.

Embi,N., Rylatt,D.B. and Cohen,P. (1979) *Eur. J. Biochem.*, **100**, 339–347.

Ferguson,M.A.J., Homans,S.W., Dwek,R.A. and Rademacher,T.W. (1988) *Science*, **239**, 753–759.

Holmes,C.F.B., Campbell,D.G., Caudwell,F.B., Aitken,A. and Cohen,P. (1986) *Eur. J. Biochem.*, **155**, 173–182.

Holt,G.D. and Hart,G.W. (1986) *J. Biol. Chem.*, **261**, 8049–8057.

Hunter,A.R. (1985) *Trends Biochem. Sci.*, **10**, 275–280.

Kennedy,L.D., Kirkman,B.R., Lomako,J., Rodriguez,I.R. and Whelan,W.J. (1985) In Berman,M.C., Gevers,W. and Opie,L.H. (eds), *Membranes and Muscle*. ICSU Press/IRL Press, Oxford, pp. 65–84.

Krisman,C.R. and Barengo,R. (1975) *Eur. J. Biochem.*, **52**, 117–123.

Mayberry,W.E., Rall,J.E. and Bertoli,D. (1965) *Biochemistry*, **12**, 2606–2611.

McCarthy,A.D. and Hardie,D.G. (1984) *Trends Biochem. Sci.*, **9**, 60–63.

Nimmo,H.G., Proud,C.G. and Cohen,P. (1976) *Eur. J. Biochem.*, **68**, 21–30.

Pitcher,J., Smythe,C., Campbell,D.G. and Cohen,P. (1987) *Eur. J. Biochem.*, **169**, 497–502.

Pitcher,J., Smythe,C. and Cohen,P. (1988) *Eur. J. Biochem.*, in press.

Poulter,L., Ang,S.-G., Williams,D.H. and Cohen,P. (1987) *Biochim. Biophys. Acta*, **929**, 296–301.

Rodriguez,I.R. and Fliesler,S.J. (1988) *Arch. Biochem. Biophys.*, **260**, 628–637.

Rodriguez,I.R. and Whelan,W.J. (1985) *Biochem. Biophys. Res. Comm.*, **132**, 829–836.

Wimmer,E. (1982) *Cell* **28**, 199–201.

Received on May 31, 1988