THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1989 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 264, No. 22, Issue of August 5, pp. 12826–12829, 1989
Printed in U.S.A.

Complete Sequences of Glucagon-like Peptide-1 from Human and Pig Small Intestine*

(Received for publication, March 23, 1989)

Cathrine Ørskov‡, Maurizio Bersani§, Anders H. Johnsen‡, Peter Højrup¶, and Jens J. Holst§

From the §Institute of Medical Physiology C, Panum Institute, University of Copenhagen, Denmark, the ‡Department of Clinical Chemistry, Rigshospitalet, University of Copenhagen, 2100 Copenhagen, Denmark and the ¶Institute of Molecular Biology, University of Odense, 5000 Odense, Denmark

In the small intestine, proglucagon is processed into the previously characterized peptide "glicentin" (proglucagon (PG) 1–69) and two smaller peptides showing about 50% homology with glucagon: glucagon-like peptide-1 and -2. It was assumed that the sites of posttranslational cleavage in the small intestine of the proglucagon precursor were determined by pairs of basic amino acid residues flanking the two peptides.

Earlier studies have shown that synthetic glucagon-like peptide-1 (GLP-1) synthesized according to the proposed structure (proglucagon 71–108 or because residue 108 is Gly, 72–107 amide) had no physiological effects, whereas a truncated form of GLP-1, corresponding to proglucagon 78–107 amide, strongly stimulated insulin secretion and depressed glucagon secretion.

To determine the amino acid sequence of the naturally occurring peptide we isolated GLP-1 from human small intestine by hydrophobic, gel permeation, and reverse-phase high performance liquid chromatography. By analysis of composition and sequence it was determined that the peptide corresponded to PG 78–107. By mass spectrometry the molecular mass was determined to be 3295, corresponding to PG 78–107 amide. Furthermore, mass spectrometry of the methyl-esterified peptide showed an increase in mass compatible with the presence of α-carboxyl amidation. Thus, the gut-derived insulinotrophic hormone GLP-1 is shown to be PG 78–107 amide.

The amino acid sequence of the human glucagon precursor was predicted by Bell and co-workers in 1983 (1) from the nucleotide sequence of a single gene encoding glucagon. The prohormone was found to consist of three glucagon-like sequences flanked by potential dibasic cleavage sites: glucagon itself and two COOH-terminally positioned amino acid sequences with approximately 50% similarity with glucagon, glucagon-like peptide-1 (GLP-1),[1] and glucagon-like peptide-2 (GLP-2) (Fig. 1). The GLP-1 sequence is well conserved between species: in all mammals studied there has been 100% identity between the GLP-1 sections of the proglucagon sequence (1–5). Even piscine GLP-1 is more than 50% similar to mammalian GLP-1 (6–8). In fact, this degree of similarity is higher than that of the glucagon sequences from the same species.

GLP-1 immunoreactivity (GLP-1-IR) has been found in extracts of both pancreas and small intestine of man (9, 10), pig (11), and rat (12), and in the pancreas of catfish (8).

By gel filtration of extracts of mammalian pancreata most workers have found that the elution position of the majority of GLP-1-IR corresponds to an apparent molecular mass of approximately 10,000 (13). This form, which has been named the major proglucagon fragment (MPGF) probably also contains the GLP-2 sequences, since it also has GLP-2 immunoreactivity, but not glucagon immunoreactivity (11).

By gel filtration of extracts of the small intestine, however, GLP-1-IR elutes at a position corresponding to a peptide with a molecular mass of approximately 3000. Judging from the predicted amino acid sequence of the prohormone and from the position of dibasic amino acids, intestinal GLP-1 would be expected to correspond to PG 72–107. In all studies reported so far, synthetic replicas of this sequence have been without any physiological effects. However, a synthetic peptide corresponding to an NH_2-terminally truncated GLP-1 (proglucagon 78–107 amide) has several interesting effects: in physiological concentrations it strongly stimulates insulin secretion, supresses glucagon secretion (14, 15), and inhibits acid secretion (16). GLP-1-IR has been shown to be released into the circulation after a mixed meal (17) and in response to a glucose load (18). Thus, GLP-1 has been suggested to be an "incretin" (19). The truncated GLP-1 could arise post-translationally if a monobasic cleavage of proglucagon occurs after Arg^{77}.

In 1988 we isolated and partly sequenced pig GLP-1 and found that the NH_2-terminal amino acid residue of pig GLP-1 corresponds to proglucagon His^{78} (14). The full sequence of porcine GLP-1 and the sequence of human GLP-1 have not previously been reported. The position of glycine as residue 108 in proglucagon, *i.e.* next to the potential COOH-terminal cleavage site (PG 109–110), suggests that the peptide could be post-translationally carboxyl-amidated (20). Recently Patzelt and Weber (21) have claimed that MPGF (which contains the GLP-1 sequence) contains one O-glycosylated serine.

In order to determine the amino acid sequence and examine the possibility of COOH-terminal amidation and O-glycosylation we have isolated GLP-1 from human and pig small intestine.

MATERIALS AND METHODS

GLP-1 Radioimmunoassay—In all purification steps GLP-1-IR was monitored by radioimmunoassay for GLP-1, using antiserum 2135

* This work was supported by the Vissing Foundation, the Danish Medical Research Council, and the Carlsberg Foundation (to P. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations used are: GLP, glucagon-like peptide; IR, immunoreactivity; PG, proglucagon; HPLC, high performance liquid chromatography; MPGF, major proglucagon fragment; PTH, phenylthiohydantoin; PDMS, plasma desorption mass spectrometry.

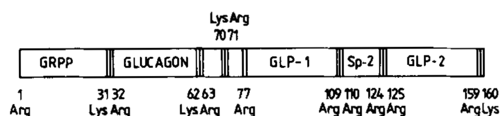

FIG. 1. **Proglucagon in mammals is a 160-residue peptide precursor.** It consists of an NH_2-terminal sequence which is similar to glicentin-related pancreatic peptide (*GRPP*) (PG 1–30), and three glucagon-like sequences: glucagon (PG 33–61), glucagon-like peptide-1 (PG 78–107 amide), and glucagon-like peptide-2 (PG 126–158). The two latter sequences are separated by an intervening sequence of 13 amino acid residues, the so-called spacer peptide-2 (*Sp-2*) (PG 111–122 amide).

raised against synthetic GLP-1, ^{125}I-labeled synthetic GLP-1, and synthetic GLP-1 (proglucagon 78–107 amide, code 7168, Peninsula Laboratories) for standards as described in Ref. 14.

Extraction Procedure—Pieces of normal human ileum were obtained during surgery on the gastrointestinal tract (approved by the Local Ethical Committee) and immediately frozen. Pieces of pig ileal mucosa were removed from anesthetized pigs (Strain LYY) and immediately frozen. Acid ethanol extracts were prepared according to method II in Ref 22. In short, frozen tissue was homogenized in 4 volumes of acid/ethanol and centrifuged. Five volumes of ice-cold diethyl ether were added to the supernatant and the aqueous phase was isolated at −50 °C. The precipitate was then dissolved in distilled water containing 1 M urea.

Purification—After extraction the peptide was purified in several steps. First, extracts of about 130 g of human small intestine and 240 g of pig small intestine, respectively, were applied to a 3 × 15-cm glass column packed with Techoprep C_{18}, 40–63 μm (HPLC Technology, Macclesfield, Cheshire, United Kingdom) and eluted with a gradient of 20–80% ethanol in water containing, in addition, 0.1% of trifluoroacetic acid (Merck). The GLP-1-IR fractions were pooled, evaporated to a smaller volume in a vacuum centrifuge (Hetovac, Heto, Hillerød, Denmark) and subjected to gel filtration on a K16/100 glass column packed with Sephadex G-50 (fine) (Pharmacia) equilibrated and eluted in 0.5 M acetic acid at a flow rate of 20 ml/h. Small amounts of ^{125}I-albumin and ^{22}NaCl were added to all samples prior to gel filtration for internal calibration. K_d, the coefficient of distribution, was then calculated for all fractions using the equation: $K_d = (V_e - V_0/V_i)$ where V_e is the elution volume of the fraction in question, V_0 the elution volume of ^{125}I-albumin, and V_i the inner volume of the column, determined as the difference between the elution volumes of ^{22}NaCl and V_0.

The GLP-1-IR fractions were pooled and all further purifications of GLP-1 were performed by reverse-phase high performance liquid chromatography (HPLC) on columns packed with Nucleosil 300, 5-μm C_4, 7-μm C_6H_5, and 5-μm C_{18} (catalog Nos. 721270, 721275, and 721273, Macherey-Nagel, Duren, Federal Republic of Germany) employing LKB HPLC equipment. The columns were eluted with a gradient of acetonitrile (grade S, Rathburn Chemicals Ltd., Walkerburn, Scotland) in water containing in addition 0.1% trifluoroacetic acid (Pierce Chemical Co.). One-min fractions were collected and subsequently subjected to analysis for GLP-1-IR. All GLP-1-IR material was collected and subjected to renewed reverse-phase HPLC, until the immunoreactive peptides seemed pure as judged by UV detection at 220 nm (Fig. 2).

HPLC Identification of Peptides—Samples containing approximately 1 nmol of: (*a*) synthetic GLP-1 (PG 78–107 amide), (*b*) purified human small intestinal GLP-1, and (*c*) synthetic GLP-1 plus purified human GLP-1 were subjected to isocratic reverse-phase HPLC on the Nucleosil 300 5-μm C_4 column. The retention times of *a*, *b*, and *c* on the column were then compared.

Amino Acid Analysis—The analysis of purified peptides were performed after hydrolysis for 20 h in 6 N HCl at 110 °C. The hydrolysates were analyzed on an Aminoquant system (Hewlett-Packard). All chemicals were analytical grade and solvents were HPLC-grade.

Sequence Analysis—This analysis of human and pig GLP-1-IR was performed on an automatic protein sequenator (Model 475A, Applied Biosystems) equipped with on-line HPLC detection of the PTH-derivatives. To improve the separation of Trp from the sequencer by-product, *N,N'*-diphenylurea, which is produced by oxidation of another by-product, *N,N'*-diphenylthiourea, a modification of the gradient (RUN 470-L) recommended by Applied Biosystems was employed.

Mass Spectrometry—Molecular mass determination was carried out on a Bio-Ion Bin 10K plasma desorption time of flight mass spectrometer (Bio-Ion AB, Uppsala, Sweden). Aproximately 20–50 pmol of peptide was applied to aluminized mylar foil coated with nitrocellulose in 2 μl of 0.1% trifluoroacetic acid and dried as described (23. After insertion of the sample in the mas spectrometer it was bombarded with fission fragments from a 10-μCi ^{252}Cf source. The spectra were recorded for 1×10^6 primary ions. The M_r of the peptide was calculated as average of the single- and double-charged molecular ions.

Methyl esterification was carried out on approximately 100 pmol of peptide (see below). Half of the sample was loaded into the mass spectrometer and spectra recorded as described above.

RESULTS

Human and pig intestinal extracts contained 57 ± 9 and 86 ± 7 pmol/g of GLP-1-IR, respectively. Three steps of the purification procedure are shown in Fig. 2.

By consecutive isocratic HPLC of (*a*) synthetic GLP-1 (PG 78–107 amide), (*b*) human small intestinal GLP-1, and (*c*) synthetic GLP-1 plus natural intestinal GLP-1, we found identical retention times with *a*, *b*, and *c* (not shown).

In Table I the results of amino acid analyses are shown. The amino acid compositions of both human and porcine GLP-1 corresponded to that of PG 78–107.

Sequence determination by Edman degradation of human and porcine GLP-1-IR, showed both peptides to be identical to PG 78–107 (Table II).

Approximately 50 pmol of the purified human GLP-1 was analyzed by plasma desorption mass spectrometry (PDMS). The resulting spectrum showed very prominent peaks at m/z 3301.9 and 1650.7, corresponding to the single and double-ionized molecular ions (Fig. 3*A*) resulting in a calculated average mass for human GLP-1 of 3300.2 ± 3.3 atomic mass

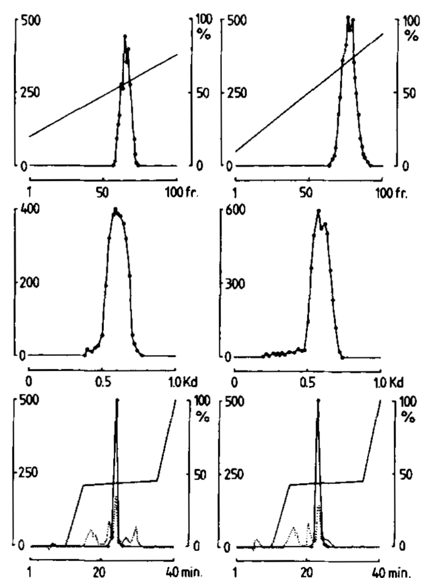

FIG. 2. **Isolation of human and porcine GLP-1.** The concentration of GLP-1-IR was monitored by radioimmunoassay (*left ordinate* scale in all six panels, nanomole/liter). The three *left panels* represent purification of human GLP-1, the three *right panels* porcine GLP-1. The two *upper panels* show the results of hydrophobic chromatography using a water-ethanol gradient. The percentage of ethanol is shown on the *right ordinate*. The two *middle panels* show gel filtration of peak fractions from the first step. The concentration of GLP-1 immunoreactivity is plotted against K_d, the coefficient of distribution, calculated as described. The two *lower panels* show the last reverse-phase HPLC (on a C_4-packed column) before sequence determination and mass spectrometry. The column was eluted with a gradient of acetonitrile in water containing 0.1% trifluoroacetic acid. The percentage of acetonitrile is shown on the *right ordinate*. *The dotted line* represents UV absorption at 226 nm.

TABLE I
Amino acid analysis of human and porcine GLP-1

For the analysis, 42 and 23 pmol, respectively, of GLP-1 were hydrolyzed. The number of residues was calculated relative to Val (2.0). Trp was not determined.

	Number of residues		
	Expected	Human GLP-1	Porcine GLP-1
Asparagine	1	1.3	1.4
Glutamine	4	4.6	4.4
Serine	3	2.8	3.2
Histidine	1	1.1	1.1
Glycine	3	3.3	4.0
Threonine	2	1.8	1.9
Alanine	4	3.7	3.8
Arginine	1	1.1	1.2
Tyrosine	1	0.9	0.9
Valine	2	2.0	2.0
Isoleucine	1	1.1	1.2
Phenylalanine	2	2.1	1.9
Leucine	2	2.5	2.2
Lysine	2	2.2	2.0

TABLE II
Sequence analysis of 250 pmol of human and 350 pmol of porcine GLP-1

Yields were corrected for background as well as lag. In both species amino acid 25 was identified, but not quantitated, partly as genuine PTH-Trp (which was only partially resolved from *N,N'*-diphenylurea) and partly as a derivative of PTH-Trp that eluted midway between *N,N'*-diphenylthiourea and *N,N'*-diphenylurea.

Amino acid No.	PTH-derivative	Human GLP-1 yield	Porcine GLP-1 yield
		pmol	
1	His	33	93
2	Ala	156	158
3	Glu	86	130
4	Gly	113	112
5	Thr	215	87
6	Phe	84	99
7	Thr	216	62
8	Ser	140	63
9	Asp	84	86
10	Val	39	73
11	Ser	24	52
12	Ser	112	51
13	Tyr	28	31
14	Leu	52	46
15	Glu	30	38
16	Gly	14	37
17	Gln	25	36
18	Ala	25	42
19	Ala	15	33
20	Lys	9	24
21	Glu	15	11
22	Phe	10	24
23	Ile	5	26
24	Ala	16	19
25	Trp	—	—
26	Leu	11	21
27	Val	5	13
28	Lys	1	8
29	Gly	5	13
30	Arg	3	7

units. This is in agreement with the theoretical value of 3297.7 atomic mass units for proglucagon 78–107 amide.

As the precision of PDMS in this mass range is insufficient to determine whether GLP-1 is COOH-terminally amidated or not, a sample was methyl-esterified by a procedure which methylates all carboxyl groups (and thus adds 14.02 atomic mass units for each) in the molecule (23). The resulting mass spectrum (Fig. 3*B*) shows the methylated average mass to be

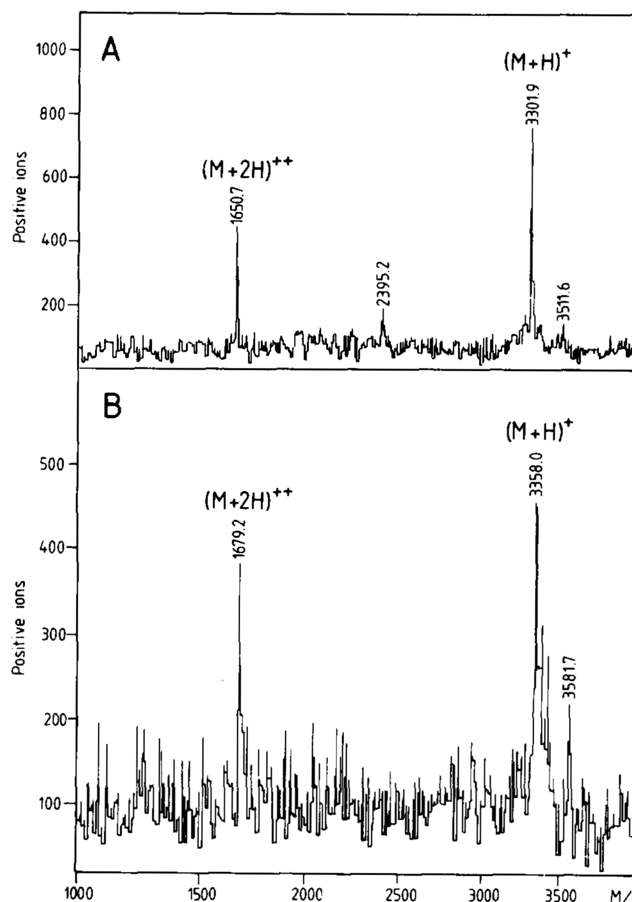

FIG. 3. *a*, PDMS spectrum of human GLP-1. Approximately 50 pmol was applied to nitrocellulose-covered aluminized mylar targets. The peak at *m/z* 2395.2 has not been identified. *b*, PDMS spectrum of methyl esterified human GLP-1. The background has been subtracted in both spectra.

3356.7 atomic mass units. This corresponds to 4.0 carboxyl groups and comparison with the sequence (containing 1 aspartic acid and 3 glutamic acid residues) shows that one carboxyl group of GLP-1 is blocked. Hence, GLP-1 is likely to be COOH-terminally amidated.

In the spectrum of amidated GLP-1 a minor peak can be observed at *m/z* 3581.7. Careful examination of the spectrum of the non-methylated species revealed a small peak at *m/z* 3511.6. The distance of these peaks to the major peaks (223.7 and 209.7 atomic mass units, respectively) shows the presence of a peptide containing one additional carboxyl group. Examination of the sequence of proglucagon showed that these masses correspond to residues PG 78–109 with a non-amidated COOH-terminal group (therorical mass 3511.9 atomic mass units).

DISCUSSION

In this work human and pig small intestinal GLP-1 were purified in sufficient amount to allow determination of their primary structures. By amino acid analysis, sequence analysis, and mass spectrometric identification it was ascertained that gut GLP-1 in pig and man is identical to PG 78–107 amide. We found no indication either by mass spectrometry or in the HPLC identification experiments of any glycosylation of GLP-1 (which would increase the molecular mass by at least 200 and might change the HPLC behavior). The lack of glycosylation of GLP-1 may indicate that glycosylation of MPGF (21) is a rat phenomenon, that another part of MPGF

other than GLP-1 is glycosylated, or that post-translational glycosylation does not take place in the small intestine as opposed to the pancreas.

By mass spectrometry we also found a small amount of GLP-1 corresponding to PG 78–109 (estimated to be less than 5% of the amount of PG 78–107 amide). The presence of PG 78–109 has not been described previously and it remains to be ascertained whether this peptide has any physiological effects and if it is an intermediate or a final processing product.

In an earlier study of human intestinal GLP-1 we found that its retention time differed slightly from that of synthetic GLP-1 (proglucagon 78–107 amide) (9). At that time we compared a small amount of synthetic peptide with the intestinal GLP-1 present in a crude extract of the small intestine. It has been shown, however, that the protein load may affect HPLC retention times (24). This finding may explain the apparent differences in retention time observed in Ref. 9.

As mentioned above, using synthetic peptides we and others have demonstrated that in contrast to PG 72–107, truncated GLP-1 (PG 78–107 amide) has several physiological effects (10, 14, 15). Our finding in this work that the major part of the naturally occurring GLP-1 in man and pig is indeed non-glycosylated PG 78–107 amide allows us to attribute these effects to the natural peptide as well. Thus GLP-1 (PG 78–107 amide) from the small intestine seems to be a new hormone with important functions in the control of metabolism and glucose turnover.

Acknowledgment—Professor T. W. Schwartz is gratefully thanked for helpful discussions.

REFERENCES

1. Bell, G. I., Sanchez-Pescador, R., Laybourn, P. J., and Najarian, R. C. (1983) *Nature* **304,** 368–371
2. Seino, S., Welsh, M., Bell, G. I., Chan, S. J.,and Steiner, D. F. (1986) *FEBS Lett.* **203,** 25–30
3. Bell, G. I., Santerre, R. F., and Mullenbach, G. T. (1983) *Nature* **302,** 716–718
4. Lopez, L. C., Frazier, M. L., Su, C.-J., Kumar, A., and Saunders, G. F. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 5485–5489
5. Heinrich, G., Gros, P., and Habener, J. F. (1984) *J. Biol. Chem.* **259,** 14082–14087
6. Plisetskaya, E., Pollock, H. G., Rouse, J. B., Hamilton, J. W., Kimmel, J. R., and Gorbman, R. (1986) *Regul. Pept.* **14,** 57–67
7. Lund, P. K., Goodman, R. H., Dee, P. C., and Habener, J. F. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 345–349
8. Andrews, P. C., and Ronner, P. (1985) *J. Biol. Chem.* **260,** 3910–3914
9. Ørskov, C., Holst, J. J., Poulsen, S. S., and Kirkegård, P. (1987) *Diabetologia* **30,** 874–881
10. George, S. K., Uttenthal, L. O., Ghiglione, M., and Bloom, S. R. (1985) *FEBS Lett.* **192,** 275–278
11. Ørskov, C., Holst, J. J., Knuhtsen, S., Baldissera, F. G. A., Poulsen, S. S., and Nielsen, O. V. (1986) *Endocrinology* **119,** 1467–1475
12. Manaka, H., Taniguchi, H., Wada, K., Takahashi, H., Takahashi, K., Katagiri, T., and Sasaki, H. (1987) *Biomed. Res.* **8,** 1–8
13. Patzelt, C., and Schiltz, E. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 5007–5011
14. Holst, J. J., Ørskov, C., Nielsen, O. V., and Schwartz, T. W. (1987) *FEBS Lett.* **211,** 169–174
15. Mojsov, S., Weir, G. C., and Habener, J. F. (1987) *J. Clin. Invest.* **79,** 616–619
16. Scholdager, B., Mortensen, P. E., Christiansen, J., Ørskov, C., and Holst, J. J. (1989) *Dig. Dis. Sci.* **34,** 703–708
17. Ørskov, C., and Holst, J. J. (1987) *Scand. J. Clin. Lab. Invest.* **47,** 165–174
18. Kreymann, B., Williams, G., Ghatei, M. A., and Bloom, S. R. (1987) *Lancet* **2,** 1300–1303
19. Creutzfeldt, W., and Ebert, R. (1985) *Diabetologia* **28,** 565–573
20. Bradbury, A. F., and Smyth, D. G. (1987) *Biosci. Rep.* **7,** 907–916
21. Patzelt, C., and Weber, B. (1986) *EMBO J.* **5,** 2103–2108
22. Newgard, C., and Holst, J. J. (1981) *Acta Endrocrinol.* **98,** 564–572
23. Nielsen, P. F., Klarskov, K., Højrup, P., and Roepstorff, P. (1988) *Biomed. Environ. Mass Spectrom.* **17,** 355–362
24. Hancock, W. S., and Sparrow, J. F. (1984) in *HPLC Analysis of Biological Compounds*, pp. 230–253, Marcel Decker, New York