Partial amino acid sequence of a human seminal plasma peptide with inhibin-like activity

N.G. Seidah, K. Ramasharma*, M.R. Sairam* and M. Chrétien+

*Laboratories of Biochemical and Molecular Neuroendocrinology and *Reproduction Research Laboratory, Clinical Research Institute of Montreal, 110 Pine Avenue West, Montreal H2W 1R7, Canada*

Received 25 November 1983; revised version received 18 December 1983

An extract of human seminal plasma was found to have inhibin-like activity. The active factor was purified to homogeneity by ion exchange chromatography, molecular sieving and high performance liquid chromatography. The purified material has a mass of approximately 5 kDa and is very basic. Amino acid analysis showed the presence of approximately 35 residues while the sequencing data allowed the determination of the N-terminal 31 amino acids. There is a possibility of an additional 2–4 residues at the C-terminus, which could not be determined.

Inhibin *Sequence* *Fragment* *Seminal plasma* *Composition*

1. INTRODUCTION

'Inhibin' was first described in 1932 [1] and was thought to have an important effect on the secretion of pituitary follicle-stimulating hormone (FSH) [2]. More recently, evidence coming from different laboratories revealed that inhibin is a polypeptide or proteinaceous molecule present in extracts of numerous gonadal-related fluids of different species [3–8]. Different types of isolation procedures have been described, but complete purification and chemical characterization of the active principle is yet to be achieved.

In human seminal plasma, different active fractions were obtained and found to have molecular masses varying from 5 to >100 kDa [9]. This heterogeneity in addition to the basic nature of the low M_r component delayed much of their characterization. The 5 kDa material was finally isolated [9,10] to homogeneity and its biological as well as chemical properties were determined. Here, we report its chemical characterization.

+ To whom correspondence should be addressed

2. MATERIALS AND METHODS

2.1. *Isolation and purification of human inhibin-like peptide*

The ethanolic precipitate obtained from human ejaculates (sperm-free) was put through a series of isolation steps including ion-exchange chromatography on sulphopropyl-Sephadex C-50, DEAE Sephadex A-25, gel filtration on Sephadex G-50, and finally successive HPLC using C18 micro Bondapak column. Details of the purification procedure and biological activities have been described [9,10]. Purity of the material was assessed by different methods including polyacrylamide gel electrophoresis at pH 4.5 which also confirmed its basic nature as suggested by its behavior during ion-exchange chromatography. Gel filtration on Sephadex-50 suggested an apparent molecular mass of 4–5 kDa.

2.2. *Amino acid analysis and sequencing procedure*

Amino acid analysis was done following acid hydrolysis (6 N HCl at 110°C for 22 h) on a Beckman model 121M analyzer. Amino acid se-

Published by Elsevier Science Publishers B.V.

quence was performed on a Beckman 890C automatic sequenator using a 0.3 M Quadrol program [11,12] and 3 mg of Polybrene (Aldrich) as a carrier. The instrument is equipped with a Sequemat P-6 autoconverter for PTH conversion (using methanolic HCl) and a model SC-510 controller for automatic program switching. The phenylthiohydantoin residues were analyzed by HPLC using an Altex 5μ-ODS column and a Varian LC-5060/Vista 401 apparatus as previously described [11,12]. The integrated yield of each PTH-amino acid was then plotted as a function of cycle number as shown in fig.1.

2.3. *Computer data bank search*

The possible homologies of the deduced sequence was verified using a computer mutation data matrix procedure as developed by the late Dr M.O. Dayhoff of the National Biomedical Research Foundation, Georgetown University in Washington, DC.

3. RESULTS

The amino acid analysis (table 1) of the isolated human seminal plasma peptide revealed that it is a highly basic peptide containing up to 35 amino acids. It is important to note the absence of Thr, Met, Cys, Pro, Leu and Tyr. The exact number of residues will be discussed later.

Table 1

Amino acid composition of inhibin-like peptide

Amino acid	Mean ± SE (n = 6)	Nearest integer	Available structure data
Lys	5.47 ± 0.29	5	5
His	6.6 ± 0.5	7	7
Arg	2.8 ± 0.06	3	3
Asp	3.8 ± 0.2	4	3
Ser	1.54 ± 0.09	1–2	1
Glu	2.5 ± 0.08	2–3	2
Gly	5.8 ± 0.4	6	5
Ala	0.96 ± 0.07	1	1
Val	2.0 ± 0.05	2	2
Ile[a]	1.0	1	1
Phe	0.9 ± 0.1	1	1

[a] Ile was taken as the reference amino acid, from which the composition was calculated

Samples were hydrolyzed in 6 N HCl at 110°C for 22 h. Thr, Met, Cys, Pro, Leu, Tyr are absent. Trp was not analyzed. Hexosamines were not detectable

The strategy employed for the sequence involved 3 consecutive sequence runs. In the first one 300 pmol of the peptide was deposited on the sequenator cup and double coupling performed for the first cycle only. This was then followed by the normal single coupling procedure. This methodology allowed the unambiguous determination of the first 20 amino acids. It was then decided to repeat the sequence with 3 nmol of the peptide using an identical strategy. In a second sequence run, due to the availability of more material at the start, it was possible to extend the sequence to 31 cycles. Although low amounts of amino acids were detected past residue 31, it was not possible to unambiguously identify them. Furthermore, at this level of starting material it became clear that the peptide preparation was contaminated by fragments of the 31 amino acid sequence determined (fig.3). These included peptides starting at residues 19, 2 and 10 of the original major sequenceable material. If we denote the 31 amino acid sequence as the N sequence (see fig.3), then the latter are then denoted N-18, N-1 and N-9, respectively. In order to confirm the sequence obtained, to possibly extend it at the carboxy-terminal and to quantitate the amounts of contaminating fragments it was then decided to repeat the sequence with the same strategy but the analysis of the residues was performed as soon as the PTH amino acid came out of the sequenator. In this case, about 30 nmol of peptide was used at the start.

In fig.1 are represented the yields of each PTH amino acid as a function of sequenator cycle number. First, it is clear that within the 31 amino acid sequence, Thr, Met, Cys, Pro, Leu, Tyr and Trp are absent, thus confirming similar results obtained with the amino acid analysis. Second, it is evident that the N sequence represents the major one and that the N-18, N-1 and N-9 fragments represent about 20, 18 and 18%, respectively, of the N peptide, as calculated from the regression lines shown in fig.2. Therefore, in order to dissect the data and calculate the yields for each one of these polypeptide chains, in fig.2 are shown the linear regression lines for the N, N-18 and N-1 se-

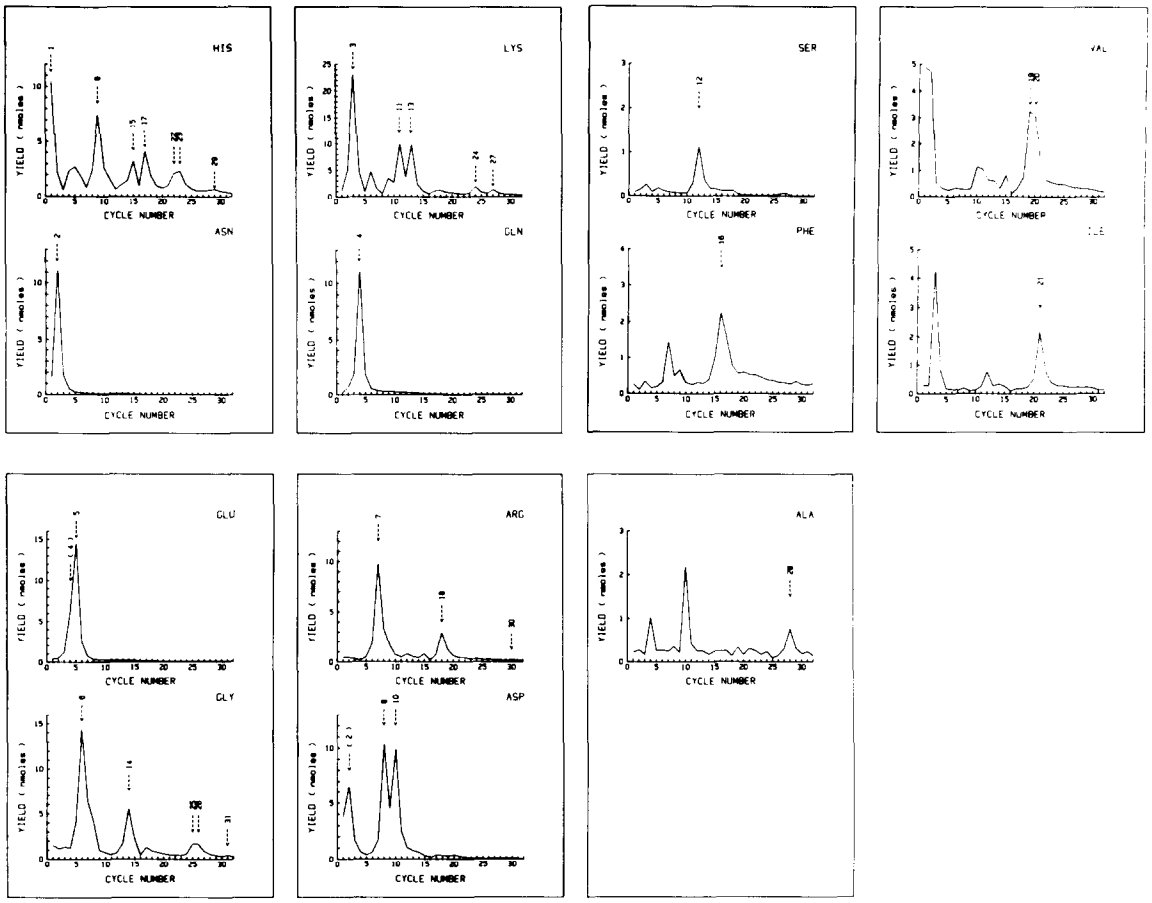

Fig.1. Yield of the various amino acids obtained during the sequence of the human inhibin-like peptide isolated. Following residue 31, no more assignment of amino acids was possible due to very low yields. The numbers above each peak represent the deduced position occupied by that amino acid along the peptide chain. Although other peaks are apparent, e.g., Ala at position 10, Ile at position 3, Phe at position 7, Ser at position 3, these are deduced to have come from the N-18 and N-9 fragments, respectively (see section 3). The contribution of the N-1 segment is apparent by the presence of a preview peak before each amino acid assigned for the N chain. Numbers in parentheses (in Asp and Glu graphics) represent the yields of either Asp or Glu originating from partial hydrolysis of the parent Asn or Gln, respectively, assigned at these positions.

quences. The regression line for the N-9 sequence could not be calculated with accuracy in view of overlapping amino acids. However, the N-9 sequence is clearly present upon analysis of the data for Ser and Phe at positions 3 and 7 of the sequence, respectively, as shown in fig.2, which would correspond to Ser 12 and Phe 16 in the N sequence (see fig.3). This strategy, therefore, allowed the determination with certainty of the sequence of the first 31 amino acids of the human seminal plasma peptide isolated. The sequence of the N-18, N-1 and N-9 peptides allows further confirmation of the validity of the assignment of the amino acids within the N sequence. Interestingly, when one carefully analyzes the data from fig.1 and 2, it becomes evident that the various chains simultaneously analyzed were sequenced with different degrees of ease. For example, although the N sequence has an initial yield of 24.4 nmol yet the repetitive yield for this chain is only 89.5%. The N-1 chain shows an initial yield of 4.4 nmol and a repetitive yield of 92.1%. As for the shorter N-18 chain, it is clear that the yield dropped sharply following the Gly–Gly sequence (residues 25–26 of

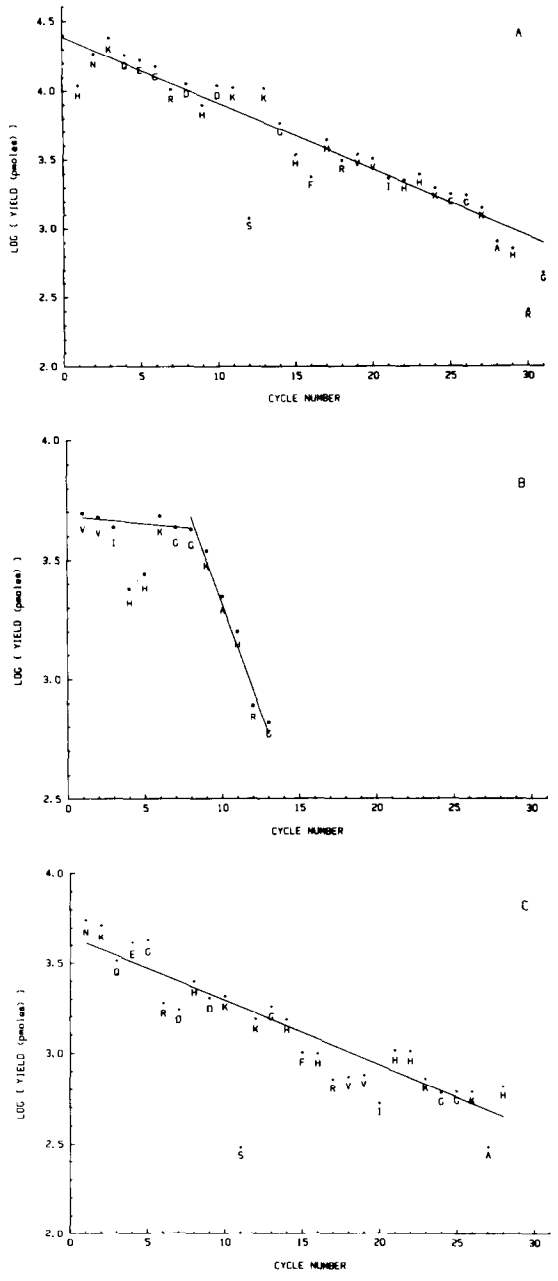

Fig.2. Calculated PTH yields of each amino acid assigned to a particular chain of the inhibin-like peptide, plotted on a semi-logarithmic scale as a function of sequenator cycle number. (A) Here the yields of each amino acid obtained for the major N sequence (see fig.3) are depicted together with the best line fit obtained by linear regression analysis. Each amino acid assigned at a given position is depicted in a one letter code. Based on the best line fit we calculate the initial yield at cycle 1 to be 24.4 nmol and the average decline in sequence yield, termed repetitive yield (obtained from the slope of the line) to be 89.5%. Some residues such as Ser are very labile, and accordingly are obtained in low yield, e.g., Ser 12. (B) Same calculations were done as for (A) for the less abundant N-18 chain. The initial yield obtained from the intercept of the best line fit with the ordinate is found to be 4.8 nmol and the repetitive yield of the first 9 cycles to be 98.5%. Following cycle 9, a sharp drop in yield is observed giving a repetitive yield for this part of the sequence as 65.9%. (C) A similar calculation for the N-1 peptide chain sequence predicts an initial yield for this peptide of 4.4 nmol and an average repetitive yield of 92.1%. However, for this chain no further sequence assignment was possible following residue 28 which could indicate that a cleavage at His 29 (of the N chain numbering) could have occurred giving rise to a 28 residues chain in the original peptide preparation. Alternatively, the sequence might have terminated precociously at that particular cycle. From the data shown in fig.1 it can be calculated, based on Phe 7 and Ser 3 that the initial yield of the N-9 chain is very similar to the N-1 chain. However, due to similarity of residues from different chains at each cycle, we cannot calculate with certainty the repetitive yield of the sequence of this N-9 fragment. From the initial yield data it can, therefore, be calculated that the original peptide preparation contained at least a mixture of 4 peptides representing the N, N-18, N-1 and N-9 segments of the proposed sequence at a ratio of 1:0.20:0.18:0.18, respectively.

the N sequence). Therefore, although the initial and repetitive yields are, respectively, 4.8 nmol and 98.5%, yet following the Gly–Gly sequence, the repetitive yield drops to 65.9%.

Accordingly, the proposed amino terminal sequence of the human seminal plasma peptide isolated and exhibiting inhibin-like activity [9,10] is depicted in fig.3.

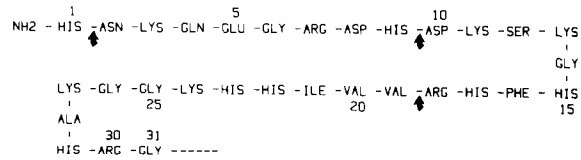

Fig.3. Proposed amino terminal sequence of the isolated human peptide exhibiting inhibin-like activity. The arrows represent the deduced cleavage sites, within the original peptide preparation, possibly responsible for the generation of the observed N-1 (His↓Arg cleavage), N-9 (His↓Asp cleavage) and N-18 (Arg↓Val cleavage) fragments.

4. DISCUSSION

From the determined amino acid composition (see table 1), it can be inferred that no more than 2–4 additional amino acids could possibly be present at the carboxy-terminus of this peptide. In fig.3 are shown the inferred cleavage positions along the peptide chain which would be responsible for the generation of the N-1, N-9 and N-18 fragments observed. Interestingly, these imply a cleavage at the carboxy-terminus of either His (for N-1 and N-9) or Arg (for N-18). It is not clear at present what the significance of such cleavage is or whether this is generated artefactually during the isolation and purification procedures employed.

In order to ascertain the possible homology between the determined sequence with that of any known protein or segment of protein, a computer sequence data bank search was initiated in collaboration with the late Dr Margaret O. Dayhoff. The results of such an exhaustive search revealed that this is a novel peptide. Furthermore, it was shown that out of 2145 protein sequences known representing a total of about 350000, 30 amino acid segments screened no more than 33% homology could be found with any known protein or segment of protein. However, such a search revealed that the best match was found to be with the segment 43–73 of either Baker's yeast enolase (10 residues out of 31) or yeast enolase (9 residues out of 31) both sequences being translated from the nucleotide sequence of yeast enolases published in [13] and [14]. Since enolase, a dimer of identical chains, catalyzes the formation of phosphoenolpyruvate in the glycolytic pathway, it is not clear at present if the observed homology has any significant meaning. Interestingly, antibodies to neuron specific enolase (NSE) and non-neuronal enolase (NNE) have been used as specific markers of neuronal and glial cells respectively [15]. These antibodies were also used to stain a number of amino precursor uptake and decarboxylation (APUD) cells [16] including those of endocrine cells. The homology of the isolated peptide with inhibin-like activity to enolase might be responsible for some cross-reactivity with these antibodies, a possibility which has to be verified experimentally. A similar type of homology was found between a recently isolated novel hypothalamic-pituitary peptide with pro-insulin [17,18]. Finally, data such as presented herein should stimulate further studies in establishing the molecular relationships among inhibin-like factors present in gonadal fluids.

ACKNOWLEDGEMENTS

This investigation was supported in part by grants from the Ford Foundation and the Medical Research Council of Canada (PG-2).

REFERENCES

[1] McCullagh, D.R. (1932) Science 76, 19–20.
[2] De Jong, F.H. and Sharpe, R.M. (1976) Nature 263, 71–72.
[3] Franchimont, P., Proyard, J.V., Hagelstein, M.T.H., Renard, C., Demoulin, A., Bourguignon, J.P. and Hustin, J. (1979) Vitam. Horm. 37, 243–302.
[4] Franchimot, P. and Channing, C.P. (1981) Intragonadal Regulation of Reproduction, Academic Press, New York.
[5] Channing, C.P. and Segal, S.J. (1982) Intraovarian Control Mechanisms, Plenum, New York.
[6] Sairam, M.R. and Madhwaraj, H.G. (1982) Int. J. Androl. 5, 205–209.
[7] Mains, S.J., Davies, R.V. and Setchell, B.P. (1979) J. Reprod. Fertil. 26, 3–14.
[8] De Jong, F.H. (1979) Mol. Cell. Endocrinol. 23, 1–10.
[9] Ramasharma, K. and Sairam, M.R. (1982) Ann. NY Acad. Sci. 383, 307–328.
[10] Ramasharma, K., Sairam, M.R., Seidah, N.G. and Chrétien, M. (1983) submitted.
[11] Seidah, N.G., Rochemont, J., Hamelin, J., Lis, M. and Chrétien, M. (1981) J. Biol. Chem. 256, 7977–7984.
[12] Lazure, C., Seidah, N.G., Chrétien, M., Lallier, R. and St. Pierre, S. (1983) Can. J. Biochem. 61, 287–292.
[13] Holland, M.J., Holland, J.P., Thill, G.P. and Jackson, K.A. (1981) J. Biol. Chem. 256, 1385–1395.
[14] Chin, C.C.Q., Brewer, J.M. and Wold, F. (1981) J. Biol. Chem. 256, 1377–1384.
[15] Schmechel, D., Marangos, P.J., Zis, A.P., Brightman, M. and Goodwin, F.K. (1978) Science 100, 313–315.
[16] Marangos, P.J., Schmechel, D., Parma, A.M., Clark, R.L. and Goodwin, F.K. (1979) J. Neurochem. 33, 313–329.
[17] Hsi, K.L., Seidah, N.G., De Serres, G. and Chrétien, M. (1982) FEBS Lett. 147, 261–266.
[18] Seidah, N.G., Hsi, K.L., De Serres, G., Rochemont, J., Hamelin, J., Antakly, T., Cantin, M. and Chrétien, M. (1983) Arch. Biochem. Biophys. 225, 525–534.