# Partial primary structure of human pregnancy zone protein: Extensive sequence homology with human $\alpha_2$ -macroglobulin

(plasma proteins/evolution/acute-phase proteins)

LARS SOTTRUP-JENSEN\*, JØRGEN FOLKERSEN<sup>†</sup>, TORSTEN KRISTENSEN<sup>‡</sup>, AND BRIAN F. TACK<sup>‡</sup>

\*Department of Molecular Biology, University of Aarhus, 8000, Aarhus C, Denmark; †Institute of Medical Microbiology, University of Odense, 5000, Odense, Denmark; and ‡Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, CA 92037

Communicated by Frank W. Putnam, August 9, 1984

ABSTRACT Human pregnancy zone protein (PZP) is a major pregnancy-associated protein. Its quaternary structure (two covalently bound 180-kDa subunits, which are further non-covalently assembled into a tetramer of 720 kDa) is similar to that of human  $\alpha_2$ -macroglobulin ( $\alpha_2$ M). Here we show, from the results of complete or partial sequence determination of a random selection of 38 tryptic peptides covering 685 residues of the subunit of PZP, that PZP and  $\alpha_2 M$  indeed are extensively homologous. In the stretches of PZP sequenced so far, the degree of identically placed residues in the two proteins is 68%, indicating a close evolutionary relationship between PZP and  $\alpha_2 M$ . Although the function of PZP in pregnancy is largely unknown, its close structural relationship to  $\alpha_2$ M suggests analogous proteinase binding properties and a potential for being taken up in cells by receptor-mediated endocytosis. In this regard our studies indicate a bait region in PZP significantly different from that present in  $\alpha_2$ M. PZP could be the human equivalent of the acute-phase  $\alpha$ -macroglobulins (e.g., rat  $\alpha_2 M$  and rabbit  $\alpha_1 M$ ) described earlier.

Human "pregnancy zone protein" (PZP)§ is one of the major pregnancy-associated plasma proteins. PZP was first described in 1959 by Smithies (2) who, upon zone-electrophoresis in starch gels, detected a characteristic band from the sera of some pregnant women. Subsequent work showed that PZP was a prominent constituent of late-pregnancy sera (3, 4). In healthy non-pregnant females and in males, PZP is present in trace amounts only (females: 10-30 mg/liter of plasma; males: <10 mg/liter of plasma) (5). During pregnancy, the plasma concentration of PZP increases and may reach levels of 1000-1400 mg/liter just before term (5-7). Small amounts of PZP have been prepared from pregnancy serum and plasma or from placental extracts by elaborate conventional procedures or more recently by immunoadsorbent techniques (8–16). PZP is a glycoprotein of  $\alpha_2$ -mobility (2-4, 8, 9) containing 10-12% carbohydrate (11-13). Preparations of PZP when subjected to denaturation under non-reducing conditions display a 360-kDa molecular species (12, 13, 16) along with high molecular mass aggregates (13). Upon reduction, 180-kDa subunits are observed (12, 13, 16), which occasionally are partially cleaved to 90-kDa fragments, as reported by Bohn and Winckler (ref. 13; also unpublished data). From gel chromatography and gradient-pore PAGE studies under non-denaturing conditions, Stimson and Farquharson (16) suggested that PZP was a tetrameric protein containing 180-kDa subunits. These features of PZP are strikingly similar to those of human  $\alpha_2$ -macroglobulin ( $\alpha_2$ M), a major proteinase binding protein containing four identical 180-kDa subunits of known primary structure (17), and indicate that PZP and  $\alpha_2$ M might be related proteins. This is further suggested by the apparent existence of common antigenic determinants in PZP and  $\alpha_2 M$  (ref. 18; unpublished data).

Proteins structurally and functionally related to human  $\alpha_2$ M have been found in the plasma of members of all major vertebrate taxa (19). Two distinct  $\alpha$ -macroglobulins ( $\alpha$ Ms) having slightly different electrophoretic mobility have been described in the rat ( $\alpha_1 M$  and  $\alpha_2 M$ ) (20–22), the rabbit ( $\alpha_1 M$ and  $\alpha_2 M$ ) (23, 24), the dog ( $\alpha_1 M$  and  $\alpha_2 M$ ) (25), and the pig ("slow"  $\alpha M$  and "fast"  $\alpha M$ ) (26). Of the two  $\alpha Ms$  found in the rat and the rabbit, rat  $\alpha_2 M$  and rabbit  $\alpha_1 M$  are acutephase reactants, since their plasma levels are greatly increased during experimental inflammation and under a variety of stress conditions. In contrast, the plasma levels of rat  $\alpha_1 M$  and rabbit  $\alpha_2 M$ , like human  $\alpha_2 M$ , are only slightly affected under these conditions (20-23, 27-32). The observation that the plasma concentration of rat  $\alpha_2 M$  is increased during pregnancy (28-30) indicates that PZP and rat  $\alpha_2 M$ could be functionally equivalent.

Here we demonstrate, from the results of sequence determination of an essentially random selection of tryptic peptides representing  $\approx 47\%$  of the sequence of the subunit of PZP, that PZP and  $\alpha_2 M$  indeed share many stretches of extremely similar sequence. This strongly suggests a close evolutionary relationship for PZP and  $\alpha_2 M$ , even closer than that already recognized for  $\alpha_2 M$  and the complement components C3 and C4 (17, 33).

# MATERIALS AND METHODS

Human PZP was isolated by immunoadsorbent techniques as described earlier (15) from pooled late-pregnancy serum. Following extensive washing with 0.1 M barbital acetate, pH 8.5/1 M NaCl, PZP was eluted from the column of Sepharose-bound anti-PZP with 3 M MgCl<sub>2</sub>/0.6 M NaCl, pH 4.8 (15). Contaminating  $\alpha_2 M$  (<5%) was removed by passing the dialyzed eluate through a column of Sepharose-bound anti- $\alpha_2 M$ . Reducing NaDodSO<sub>4</sub>/PAGE showed that this preparation, which was used for the tryptic digest described below, contained about 30% 90-kDa cleavage products.

Hydrolysates of proteins and peptides were analyzed on a Beckman 6300 instrument. Automated sequenced analyses on 1- to 5-nmol samples of proteins and peptides were performed on a Beckman 890D Sequencer in the presence of pretreated Polybrene (34) using the 0.1 M Quadrol program supplied by the manufacturer. Phenylthiohydantoin derivatives of amino acids were analyzed by HPLC on a 4  $\times$  250 mm column packed with DuPont Zorbax C<sub>18</sub> material in a Hewlett-Packard 1084B instrument equipped with a fixed

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PZP, pregnancy zone protein (human);  $\alpha M$ ,  $\alpha$ -macroglobulin; C, complement.

<sup>&</sup>lt;sup>§</sup>The pregnancy zone protein (PZP) is also known as pregnancy-associated  $\alpha_2$ -glycoprotein, pregnancy-associated globulin,  $\alpha_2$ -acutephase glycoprotein,  $\alpha_2$ -pregnoglobulin, pregnancy-associated  $\alpha$ macroglobulin, schwangerschafts protein-3, Xh antigen, and Pa-1 (1).

wavelength detector (254 nm) essentially as described (35). Reagents and solvents for sequence analysis were purchased from Beckman and from Burdick and Jackson (Muskegon, MI).

## **EXPERIMENTAL**

PZP (50 nmol of 180-kDa subunit) containing no detectable impurities was reduced with dithiothreitol and alkylated with ['H]ICH<sub>2</sub>COOH essentially as described for the sequence analysis of  $\alpha_2 M$  (34). After gel chromatography on a 2.5  $\times$ 100 cm column of Sephadex G-25 F (Pharmacia), the void volume fraction (50 ml) was digested with 0.2 mg of N-tosylphenylalanyl chloromethyl ketone-treated bovine trypsin (Worthington) for 2 hr at 22°C. The digestion was stopped by the addition of 0.2 mg of soybean trypsin inhibitor (Worthington) and the solution was freeze dried. The tryptic peptides were resolved into 29 pools by DEAE-Sephacel chromatography in NH<sub>4</sub>HCO<sub>3</sub>-containing buffers (34) (Fig. 1). Following freeze drying, the peptides in selected pools were redissolved in 0.5 ml of a 1% aqueous trifluoroacetic acid solution (Pierce) and further separated by HPLC on a 8  $\times$ 250 mm column of Vydac C<sub>4</sub> material (300-nm pore size). The peptides were eluted at 43°C with a linear gradient of

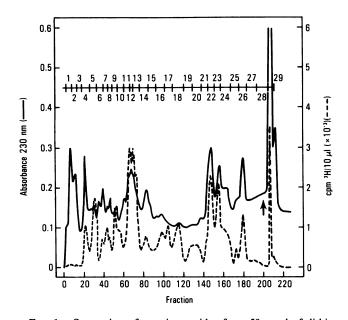


FIG. 1. Separation of tryptic peptides from 50 nmol of dithiothreitol-reduced and [<sup>3</sup>H]carboxymethylated PZP on a 0.75  $\times$  22 cm column of DEAE-Sephacel. The column was equilibrated with 10 mM NH<sub>4</sub>HCO<sub>3</sub> and eluted with gradients of NH<sub>4</sub>HCO<sub>3</sub>: from 10 mM to 250 mM (100 ml + 100 ml) and from 250 mM to 500 mM (50 ml + 50 ml). The column was finally eluted with 0.25 M NH<sub>4</sub>HCO<sub>3</sub>/3 M guanidine  $\cdot$  HCl (  $\uparrow$  ). The flow rate was 6 ml/hr and 1.5-ml fractions were collected. The separation was monitored by measuring the absorbance at 230 nm and by determining the amount of radioactivity in 10- $\mu$ l aliquots from each fraction. Except for pool 29, which was first resalted into 0.1 M NH<sub>4</sub>HCO<sub>3</sub> by Sephadex G-25 gel chromatography, all pools were directly freeze dried. Selected pools were further separated by reverse-phase HPLC (results not shown). Pools 1-3, 6, 21, 25, 27, and 28 were not investigated. No pure peptides were recovered from pools 14, 15, 19, 20, and 29. The stretches of sequence from PZP shown in Fig. 2 were obtained from the following pools: pool 4: 71-85, 86-92, 129-136, 137-149, 285-296, 317-326, and 338-352; pool 5: 782-788, 1125-1129, and 1298-1306; pools 7 and 8: 23-45, 160-181, 485-498, 499-516, 665-687, 892-917, and 1168-1187; pools 9 and 10: 192-203, 215-235, 242-251, 1142-1156, 1307-1321, and 1322-1333; pools 11 and 12: 428-461, 997-1008, and 1100-1113; pool 13: 49-70, 303-316, and 744-764; pools 16 and 17: 395-427, 807-818, and 879-891; pool 18: 518-553; pool 22: 252-274 and 710-743; pool 23: 819-830; pool 24: 603-647; pool 26: 1426-1451. The numbering of stretches in PZP is that of  $\alpha_2 M$ .

CH<sub>3</sub>CN [from 10% CH<sub>3</sub>CN/90% 0.1% trifluoroacetic acid to 70% CH<sub>3</sub>CN/30% 0.1% trifluoroacetic acid (36)] at a flow rate of 2 ml/min. All separations were performed with a Waters instrument equipped with a 214-nm fixed wavelength detector.

### RESULTS

The amino acid composition of PZP, prepared as in ref. 15, is given in Table 1 and is in fair agreement with that determined earlier by Von Schoultz and Stigbrand (12). The composition of PZP is remarkably similar to that of human  $\alpha_2 M$  (17). Although the preparation of PZP contained  $\approx 30\%$  cleaved subunits, the sequence of the first 19 residues could be confidently identified by using "background" subtraction as in ref. 37 (Fig. 2). The NH<sub>2</sub>-terminal sequences of PZP and human  $\alpha_2 M$  are strongly homologous (the sequence of PZP starts at position 4 in human  $\alpha_2 M$ ). The extent of homology in this region between human  $\alpha_2 M$  and PZP is about the same as for human  $\alpha_2 M$  and rat  $\alpha_2 M$  (17).

To investigate if PZP and  $\alpha_2M$  would share extended stretches of homologous sequence, a tryptic digest of reduced and [<sup>3</sup>H]carboxymethylated PZP was investigated. Although such a digest would be exceedingly complex and a number of larger hydrophobic peptides were likely to be recovered in a low yield (34), an essentially random selection of relatively small peptides would be obtained. The tryptic peptides from PZP were initially fractionated by DEAE-Sephacel ion-exchange chromatography (Fig. 1) followed by final purification by reverse-phase HPLC (data now shown).

The results of complete or partial sequence determination of 38 pure tryptic peptides, accounting for 685 residues, are given in Fig. 2. In almost every case, even for the short peptides, the sequences obtained from PZP could be readily aligned with corresponding regions in the human  $\alpha_2 M$  sequence shown in Fig. 2 (data from ref. 17). In a few instances small deletions had to be introduced in either sequence to obtain maximal alignment of the sequences. These areas (residues 256–260, 401–407, and 618–630) probably constitute parts of the surface regions in the two proteins. One major 28-residue tryptic peptide did not contain arginyl or lysyl residues and probably represents the COOH-terminal tryptic peptide of PZP. This peptide matches well with  $\alpha_2 M$ 

Table 1. Amino acid composition of PZP and  $\alpha_2 M$ 

			-
A	mino acid	PZP*	$\alpha_2 M^{\dagger}$
	Asx	131	115
	Thr	96	100
	Ser	129	122
	Glx	189	188
	Pro	78	76
	Gly	87	90
	Ala	99	92
	Cys	20	24
	Val	118	137
	Met	25	25
	Ile	61	57
	Leu	141	134
	Tyr	55	56
	Phe	65	62
	Lys	76	87
	His	31	39
	Trp	ND	11
	Arg	54	46

ND, not determined.

\*The result of duplicate analyses of the material used for the tryptic digest of PZP (6 M HCl/0.1% phenol; 110°C, 20 hr). No correction has been made for hydrolytic loss or incomplete hydrolysis. The composition is calculated on the basis of 54 residues of arginine. <sup>†</sup>Calculated from the complete sequence of  $\alpha_2 M$  (17).

e2M PZP	100 S ¥ S G K P Q Y M Y L Y P S L L H T E T T E K G C Y L L S Y L N E T Y T Y S A S L E S Y R G N R S L F T D L E A E N D Y L H C Y A F A Y P K S S S N E E Y M F L T Y Q Y K G P T Q E F K K R T T Y M Y K T E <u>P Q Y M Y L Y P S L L H T E</u> A Y K/I <u>G C Y L L S</u> H <u>L</u> X <u>E T Y T Y S A S L E</u> E G <u>R</u> / / / <u>S L F T D L Y A E K D</u> L F <u>H C Y S F</u> T L <u>P</u> R/I <u>S</u> A S S <u>E Y</u> A <u>F L</u> S I <u>Q</u> I <u>K G P T Q</u> D <u>F</u> R/
a2M PZP	200 NEDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLEGGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEHPFTVEE / <u>DENFRP</u> R/ <u>NELIPLIY</u> LEN <u>P</u> R///L <u>E</u> A <u>G</u> IN <u>QLSFPLSSEP</u> I <u>QGSY</u> R///IQ <u>HPFTVEE</u>
azM PZP	300 F V L P K F E V Q V T V P K I I T I L E E E M N V S V C G L Y T Y G K P V P G H V T V S I C R K Y S D A SD C H G E D S Q A F C E K F S G Q L N S H G C F Y Q Q V K T K V F Q L K R K E Y E M K L H T E F V L/ /V L N <u>C</u> D K Q E V <u>C E E F S Q Q L N S</u> K////M L <u>Q</u> I T N T G F <u>E M K</u> /
e2M P ZP	400 AQIQEEGTVVELTGRQSSEITRTITKLSFVKVDSHFRQGIPFFGQVRLVDGKGVPIPNKVIFIRGNEANYYSNATTDEHGLVQFSINTTNVMGTSLTVRV /IR <u>EEGT</u> DLEVTANR/I <u>SEIT</u> NIVSK/// <u>QGIPFFAQV</u> LLVDGK///SFTVHV /SFTVHV
a2M PZP	500 MYKDRSPCYGYQNYSEEHEEAHHTAYLYFSPSKSFYHLEPMSHELPCGHTQTYQAHYILNGGTLLGLKKLSFYYLIMAKGGIYRTGTHGLLYKQEOMKOE MLCFHYS <mark>HYAEEHQGA</mark> Q <u>HTANR/YFS</u> LSG <u>S</u> YI <u>HLEP</u> YAGT <u>LPCGHT</u> ETIT <u>AHY</u> T <u>LNK/QAMGQLSELSFHYL</u> T <u>MAK///SGTH</u> TLPYESG <u>DNK/G</u> S
e2M PZP	<sup>600</sup> F S I S I P V K S D I A P V A R L I Y A V L P T G D V I G D S A K Y D V E N E L A N K V D L S F S P S Q S L P A S H A H L R V T A A P Q S V C A L R A V D Q S V L L M K P D A E L S A S S V Y N L L P <u>F</u> A L <u>S F P Y E S D V A P I A R</u> /X F <u>I F A I L P D G E Y Y G D S</u> E K F E I <u>E N</u> C <u>L A N K V D L</u> H <u>F</u> X <u>P</u> A <u>Q</u> /
a2M PZP	<sup>698</sup> / <u>D_L</u> TN <u>FP</u> DN V DQ D D E D C I N R H N V Y I N G I T Y T P V S S T N E K D M Y S F L E D M G L K A F T N S K I R K P K M C P Q L Q Y E M H G P E G L R V G F Y E S D V M G R G H A R L V / <u>D_L</u> TN <u>FP</u> D N V D Q Q E E <u>E</u> Q G H <u>C</u> P R P F F I H N T A <u>1</u> - <u>Y</u> Y P L <u>S S</u> - <u>M E A D</u> I <u>Y S F L</u> / /S <u>C</u> S V T P S V S A - <u>G</u> A V G Q <u>G</u> Y <u>Y</u> G A/
e2M PZP	798 H VEEPH TET V R K Y F P ET M I M D L V V V N S A G V A E V G V T V P D T I T E H K A G A F C L S E D A G L G I S S T A S L R A F Q P F F V E L T M P Y S V I R G E A F T L K A T V L N Y L P K C /S Y F P ET M I M E L V A Y Q S S G V A E V G V T V P D T I T E H K /A G A F C L S E D A G L G I S S T A S L R / /G E V F T L K/
e2M PZP	888 I R V S V Q L E A S P A F L A V P V E K E Q A P H C I C A N G R N T V S W A V T P K S L G N V N F T V S A E A L E S Q E L C G T E V P S V P E H G R K D T V I K P L L V E P E G L E K E T T F N S L C / <u>A S P A F L A</u> S Q N T <u>K</u> /G E E S Y <u>C I C</u> G <u>N</u> E <u>R</u> / / T <u>L L V E</u> A <u>E G</u> I <u>E</u> Q <u>E</u> K/ <u>T F</u> S <u>S</u> M T <u>C</u>
e2M PZP	998 PSGGEVSEELSLKLPPNVVEESARASVSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEIKSKAIGYLNTGYQRQLNYH A <u>SG</u> A <u>EVSE</u> Q <u>LSLKLPPNVV</u> / <u>H</u> Q
a2M PZP	1098 D G S Y S T F G E R Y G R N Q G N T W L T A F V L K T F A Q A R A Y I F I D E A H I T Q A L I W L S Q R Q K D N G C F R S S G S L L N N A I K G G V E D E V T L S A Y I T I A L L E I P L T V T H P V V <u>D G S Y S T F G E R</u> /
a2M PZP	1198 R NALF CLE SAWK TAQE G DHG SHVYTKALLAYAFALAG NQDK R KE VLK SL NE EAVKK DN SVHWE R PQK PKAP VGH FYE PQAP SAE VEMT SYVLLAYLTAQP / <u>NALF CLE SAW</u> NY <u>A</u> K/// <u>ALLAY</u> // <u>E</u> T <u>L</u> N <u>SL</u> DK <u>EAVK</u> E <u>DN</u> L/// <u>APVGH</u> L <u>Y</u> Q <u>PQAP SAE VEMT S</u> /
n2M PZP	1298 A P T S E D L T S A T H I V K H I T K Q Q H A Q G G F S S T Q D T V V A L H A L S K Y G A A T F T R T G K A A Q V T I Q S S G T F S S K F Q V D H H N R L L L Q Q V S L P E L P G E Y S M K V T G E G C / <u>C</u>
a2M PZP	1398 VYLQTSLKYNILPEKEEFPFALGVQTLPQTCDEPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSNHVLIYLDK <u>VYLQTS</u> MK <u>YNILPEKE</u> DS <u>PFALK/YQT</u> Y <u>PQTCD</u> GH <u>K</u> /
azM PZP	1451 V S N Q T L S L F F T V L Q D V P V R D L K P A I V K V Y D Y Y E T D E F A I A E Y N A P C S K D L – – G N A / <u>V Y D Y Y E T D E</u> S V <u>A E Y</u> I <u>A P C S</u> S <u>D</u> T E H <u>G N</u> V

FIG. 2. Alignment of completely or partially sequenced tryptic peptides from PZP with corresponding stretches from  $\alpha_2 M$ . The complete sequence of human  $\alpha_2 M$  is from ref. 17. The amino acid sequences are given in the single-letter code (38). X denotes an unidentified amino acid residue. Identically placed residues in the two proteins are underlined. The first 19 residues of PZP were determined by sequenator analysis of the intact preparation. In addition to a major threonine residue, minor amounts of asparagine, serine, and glycine were also observed in step 1.

and suggests a 2-residue deletion at position 1448-1449 in the  $\alpha_2 M$  sequence. In some regions the sequences of longer stretches of PZP and  $\alpha_2 M$  are nearly identical (e.g., the stretches corresponding to residues 5-41 and 710-764 of  $\alpha_2$ M). On the average the stretches determined contain 68% identically placed residues. About one-half of the dissimilar residues involve chemically similar residues or exchanges that retain a charged or hydrophilic residue at a given position. Of the expected 20-24 half-cystine residues in the PZP monomer, 18 have been determined, including the half-cystine residue that by analogy with  $\alpha_2 M$  (17), presumably bridges two-subunits to form a 360-kDa half molecule. As seen from Fig. 3, the disulfide bridge pattern of PZP is likely to be very similar to that of  $\alpha_2 M$ . The half-cystine residue at position 540 has no counterpart in  $\alpha_2 M$ . Since no peptides have been recovered from the region 923–996, which in  $\alpha_2 M$ is largely constituted by a 56-residue tryptic peptide containing the  $\beta$ -cysteinyl- $\gamma$ -glutamyl thiol ester structure (17, 39), it is not yet apparent from the sequence data if PZP also contains such a structure. When the PZP preparation used in these studies was examined for its ability to incorporate  $[^{3}H]CH_{3}NH_{2}$  covalently, <0.02 mol of CH<sub>3</sub>NH<sub>2</sub> per mol of PZP subunit was bound. Of course, the method of purification of PZP might be expected to inactivate its putative proteinase-inhibiting activity and cleave any internal thiol ester bonds.

The partial sequence of a major half-cystine-containing peptide showed only a modest degree of homology with  $\alpha_2 M$ and was tentatively located at position 665–687. The region 681–686 in  $\alpha_2 M$  contains the primary sites of proteolytic cleavage and has been referred to as the bait region (17). The results of NH<sub>2</sub>-terminal sequence determination of the present PZP preparation, which contained 30% 90-kDa cleavage products, did not allow the identification of the sequences at the cleavage sites. Further evidence in favor of a bait region in PZP was obtained from an experiment in which PZP was treated with a 2-fold molar excess of trypsin for 2 min. Upon reducing NaDodSO<sub>4</sub>/PAGE, the complete disappearance of the 180-kDa subunit was accompanied by a corresponding increase in the amount of 90-kDa fragments (results not shown).

No tryptic peptides were recovered from the long stretches 1168–1297 and 1334–1425. During the sequence determination of  $\alpha_2$ M the recovery of tryptic peptides from these regions was low due to their large size and their tendency to form aggregates (34). As seen in Fig. 1, a set of strongly retained peptides could be eluted from the DEAE-Sephacel column by 3 M guanidine·HCl/0.25 M NH<sub>4</sub>HCO<sub>3</sub>. However, none of these peptides could be recovered in appreciable yield after an attempt of purification by reverse-phase HPLC.

### DISCUSSION

The results of NH<sub>2</sub>-terminal sequence determination and partial or complete sequence determination of tryptic peptides from highly purified PZP show that this human plasma protein is a close homolog of human  $\alpha_2 M$ . The 38 peptides presented in Fig. 2 account for 685 residues of the  $\approx 1450$ residues of the subunit of PZP. Sixty-eight percent of the residues in these peptides are identical to those of  $\alpha_2 M$ . Although no peptides were recovered from certain areas of PZP, notably the COOH-terminal half, it is likely that the overall degree of sequence homology between PZP and  $\alpha_2 M$ will turn out to be very high. Including chemically similar residues (D = E, T = S, V = M = I = L, F = Y = H = W, and K = R), the extent of homology increases to about 76%, indicating that the subunits of PZP and  $\alpha_2 M$  have essentially the same tertiary structure. Both proteins appear to have the

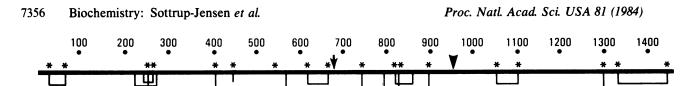


FIG. 3. Schematic representation of the 180-kDa subunit of  $\alpha_2 M$ , emphasizing the disulfide bridge pattern (data from ref. 17). Half-cystine residues found in PZP are labeled by an asterisk (\*). The activation cleavage area ( $\mathbf{\Psi}$ ) and the  $\beta$ -cysteinyl- $\gamma$ -glutamyl thiol ester site ( $\mathbf{\nabla}$ ) of  $\alpha_2 M$ are indicated.

same quaternary structure (identical 180-kDa units pair-wise disulfide bridged and further assembled to form tetramers of 720 kDa) (16, 17).

It has been pointed out recently that the proteins human  $\alpha_2$ M, murine C3, and human or murine C4, which all contain an internal reactive  $\beta$ -cysteinyl- $\gamma$ -glutamyl thiol ester (40– 43), display large segments of homologous sequence, indicating a common evolutionary origin (33). However, the homologous regions of  $\alpha_2$ M and C3, which account for 75% and 67% of their sequence, respectively, only contain 19-31% identically placed residues (33). The mutual sequence identity between C3 (44, 45), C4 (46), and an  $\approx$ 491-residue segment of C5, spanning the  $\beta$ -chain- $\alpha$ -chain junction (47) is 35-40%, indicating that these proteins, which all circulate as multiple-chain 190- to 200-kDa plasma proteins, could constitute one subset of a group of proteins related to  $\alpha_2 M$ , whereas  $\alpha_2 M$ , PZP,  $\alpha Ms$  from different animal species, and perhaps the proteinase binding protein from chicken egg white, ovostatin (48), could constitute another subset. This subset is characterized as tetrameric proteins containing subunits of  $\approx 180$  kDa, which may or may not be proteolytically processed. Interestingly, both ovostatin (48) and C5 (49) apparently do not contain an internal thiol ester.

Preparations of PZP have immunosuppressive properties under certain experimental conditions (e.g., by inhibiting the phytohemagglutinin and concanavalin A stimulation of T lymphocytes) (50, 51). These properties are thought to be of importance in the maintenance of the immunologically privileged status of the fetus (50-52). Similar phenomena have been reported for  $\alpha_2 M$  (53), and it has been suggested that these properties might reside in peptide(s) associated with or liberated from  $\alpha_2 M$  during  $\alpha_2 M$ -proteinase complex formation (54). It is not yet clear whether the effects observed with  $\alpha_2$ M preparations are due to  $\alpha_2$ M or are due to a slight contamination with PZP. Most likely, PZP, by virtue of its close similarity with  $\alpha_2 M$ , has been present as a contaminant in most  $\alpha_2 M$  preparations utilized so far.

In contrast to  $\alpha_2 M$  (31, 32, 55), highly elevated plasma levels of PZP are found during pregnancy. A less-pronounced increase in the plasma level of PZP is found in females using oral contraceptives, in males who receive estrogens during treatment for prostatic cancer, and in a variety of surgical and malignant conditions (56-59). These features indicate that PZP could be the human equivalent of the acute-phase protein rat  $\alpha_2 M$  that is also elevated in pregnancy (28-30). Like human  $\alpha_2 M$ , the plasma concentration of rat  $\alpha_1 M$  is relatively stable during the acute-phase response (20-22). For both rat  $\alpha$ Ms, the liver is probably the major site of synthesis (60-62) as it is for  $\alpha_2 M$  (unpublished data). Although PZP has been localized in the villous parenchymal tissue of the placenta and in trophoblasts (63), this organ is probably not the site of its synthesis (64). Though in vitro estrogendependent synthesis of PZP has been demonstrated in cultures of human leukocytes (65-67), synthesis of PZP by the liver hepatocytes is probably quantitatively most important (68, 69).

The demonstration here that human plasma contains a tetrameric protein, PZP, which is strongly related to  $\alpha_2 M$  in terms of primary structure, subunit organization (12, 13, 16), and potential to become specifically cleaved near the middle of its subunits, as also recognized earlier (13), suggests that the function of PZP, although still largely obscure, might be analogous to that of  $\alpha_2 M$ . In particular, investigations on the ability of PZP to form complexes with proteinases and the possible specific clearance of these complexes by receptormediated endocytosis will be important for defining the role of PZP in pregnancy. Furthermore, the study of the gene structures of PZP and  $\alpha_2 M$  may contribute to a deeper understanding of the structural elements that control the expression of these two proteins.

Part of this work was carried out while L.S.-J. was a Visiting Investigator at the Department of Immunology, Scripps Clinic and Research Foundation. This work was supported by the Danish Science Research Council, by the NATO Science foundation, by the Danish Medical Science Research Council, and by U.S. Public Health Service Grant AI 19222 from the National Institutes of Health. This is publication no. 3521 IMM from the Research Institute of Scripps Clinic.

- 1. Berne, B. H., Bohn, H., Hofmann, R., Klausch, B., Straube, W., Horne, C. H. W., Kasukawa, R., Rittner, C., Stimson, W. H. & Than, G. N. (1975) Lancet i, 367-368.
- Smithies, O. (1959) Adv. Protein. Chem. 14, 65-133.
- Cooper, D. W. (1963) Nature (London) 200, 892. 3.
- DeAlvarez, R. R. & Thompson, I. E. (1964) Obstet. Gynecol. 4. 23, 640–645.
- Folkersen, J., Teisner, B., Grunnet, N., Grudzinskas, J. G., 5. Westergaard, J. G. & Hindersson, P. (1981) Clin. Chim. Acta 110, 139-145.
- Von Schoultz, B. (1974) Am. J. Obstet. Gynecol. 119, 792-797.
- Than, G. N., Csaba, I. F., Szabo, D. G., Karg, N. J. & Novak, P. F. (1976) Vox Sang. 30, 134-138.
- 8.
- Kueppers, F. (1969) Humangenetik 7, 98-103. Bohn, H. (1971) Arch. Gynäk. 210, 440-457.
- Straube, W., Klausch, B., Hofmann, R. & Brock, J. (1972) 10. Arch. Gynäk. 212, 230-245.
- Stimson, W. H. & Eubank-Scott, L. (1972) FEBS Lett. 23, 11. 298-302.
- Von Schoultz, B. & Stigbrand, T. (1974) Biochim. Biophys. 12. Acta 359, 303-310.
- Bohn, H. & Winckler, W. (1976) Blut 33, 377-388. 13.
- Stigbrand, T., Damber, M.-G. & Von Schoultz, B. (1978) Acta 14. Chem. Scand. B32, 717-719.
- Folkersen, J., Teisner, B., Ahrons, S. & Svehag, S.-E. (1978) 15. J. Immunol. Methods 23, 117-125.
- Stimson, W. H. & Farquharson, D. M. (1978) Int. J. Biochem. 16. 9, 839-843.
- Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Wierz-17. bicki, D. M., Jones, C. M., Lønblad, P. B., Magnusson, S. & Petersen, T. E. (1984) J. Biol. Chem. 259, 8318-8327
- Weström, B. R., Karlsson, B. W. & Ohlsson, K. (1982) Hop-18. pe-Seyler's Z. Physiol. Chem. 364, 375-381.
- Starkey, P. M. & Barrett, A. J. (1982) Biochem. J. 205, 91-95. 19.
- Ganrot, K. (1973) Biochim. Biophys. Acta 295, 245-251. 20.
- Gordon, A. H. (1976) Biochem. J. 159, 643-650. 21.
- Nelles, L. P. & Schnebli, H. P. (1982) Hoppe-Seyler's Z. 22. Physiol. Chem. 363, 677-682.
- 23. Lebreton de Vonne, T. & Mouray, M. H. (1968) C.R. Acad. Sci. (Paris) Ser. D. 266, 1076-1079.
- Bloth, B., Chesebro, B. & Svehag, S.-E. (1968) J. Exp. Med. 127, 749-756. 24.
- 25. Ohlsson, K. (1971) Biochim. Biophys. Acta 236, 84-91.
- Weström, B. (1979) Hoppe-Seyler's Z. Physiol. Chem. 360, 26. 1869-1878
- 27. Morelis, P., Ambrosioni, J.-C., Got, R. & Fontanges (1969) C.R. Acad. Sci. (Paris) Ser. D. 269, 1453-1454.

- 28. Beaton, G. H., Selby, A. E., Veen, M. J. & Wright, A. M. (1961) J. Biol. Chem. 236, 2005-2008.
- 29 Heim, W. G. (1962) Nature (London) 193, 491.
- Weimer, H. E. & Benjamin, D. C. (1965) Am. J. Physiol. 209, 30. 736-744.
- Wilding, P., Adham, N. F., Mehl, J. W. & Haverback, B. J. 31. (1967) Nature (London) 214, 1226-1227.
- Ganrot, P. O. & Bjerre, B. (1967) Acta Obstet. Gynecol. 32. Scand. 46, 126-137.
- Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Lønblad, 33. P. B., Jones, C. M., Wierzbicki, D. M., Magnusson, S., Domdey, H., Wetsel, R. A., Lundwall, Å., Tack, B. F. & Fey, G. H. (1985) Proc. Natl. Acad. Sci. USA 82, in press.
- 34. Sottrup-Jensen, L., Stepanik, T. M., Jones, C. M., Lønblad, P. B., Kristensen, T. & Wierzbicki, D. M. (1984) J. Biol. Chem. 259, 8293-8303.
- Zimmerman, C. L. & Pisano, J. J. (1977) Methods Enzymol. 35. 47, 45-51.
- Mahoney, W. C. & Hermodson, M. A. (1980) J. Biol. Chem. 36. 255, 11199–11203.
- 37. Mortensen, S. B., Sottrup-Jensen, L., Hansen, H. F., Petersen, T. E. & Magnusson, S. (1981) FEBS Lett. 135, 295-300.
- 38. IUPAC-IUB Commission on Biochemical Nomenclature (1968) J. Biol. Chem. 243, 3557-3559.
- 39. Swenson, R. P. & Howard, J. B. (1980) J. Biol. Chem. 255, 8087-8091.
- Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L. & Prahl, J. W. (1980) Proc. Natl. Acad. Sci. USA 77, 5764-5768.
- 41. Sottrup-Jensen, L., Petersen, T. E. & Magnusson, S. (1980) FEBS Lett. 121, 275-279.
- 42. Janatova, J. & Tack, B. F. (1981) Biochemistry 20, 2394-2404.
- Campbell, R. D., Gagnon, J. & Porter, R. R. (1981) Biochem. 43. 7. 199, 359–370.
- 44. Lundwall, Å., Wetsel, R. A., Domdey, H., Tack, B. F. & Fey, G. H. (1984) J. Biol. Chem., in press. Wetsel, R. A., Lundwall, Å., Davidson, F. F., Tack, B. F.,
- 45 Gibson, T. & Fey, G. H. (1984) J. Biol. Chem., in press.
- Belt, K. T., Carroll, M. C. & Porter, R. R. (1984) Cell 36, 907-46. 914.
- Lundwall, Å., Wetsel, R. A., Kristensen, T., Whitehead, A. S., Woods, D., Ogden, R. C., Colten, H. R. & Tack, B. F. 47. (1985) J. Biol. Chem., in press.
- Nagase, H., Harris, E. D., Woessner, F. J. & Brew, K. (1983) 48 J. Biol. Chem. 258, 7481-7489.

- 49. DiScipio, R. G., Smith, C. A., Müller-Eberhard, H. J. & Hugli, T. E. (1983) J. Biol. Chem. 258, 10629-10636.
- Von Schoutz, B., Stigbrand, T. & Tärnvik, A. (1973) FEBS 50. Lett. 38, 23-26.
- 51. Stimson, W. H. (1976) Clin. Exp. Immunol. 25, 199-206.
- Stimson, W. H. (1972) Lancet i, 684. 52.
- 53. Hubbard, W. H., Hess, A. D., Hsia, S. & Amos, D. B. (1981) J. Immunol. 126, 292-299.
- 54. Hubbard, W. H., Anderson, B. D., Balber, A. E., Proud, G. M., Alomran, A. J. & Shenton, B. K. (1983) Ann. N.Y. Acad. Sci. 421, 332-339.
- 55. Horne, C. H. W., Howie, P. W., Weir, P. J. & Goudie, R. B. (1970) Lancet i, 49-50.
- Bohn, H. (1973) Blut 26, 205-209. 56.
- Horne, C. H. W., McLay, A. L. C., Tavadia, H. B., Carmi-chael, I., Mallingson, A. C., Yeung Laiwah, A. A. C., Thom-57. as, M. A. & McSween, R. N. M. (1973) Clin. Exp. Immunol. 13, 603-611.
- 58. Berne, B. H. (1974) Fed. Proc. Fed. Am. Soc. Exp. Biol. 33, 290.
- 59. Horne, C. H. W., Bohn, H., McLay, A. L. C., Wood, E. H. & Thomson, W. D. (1975) Behring Inst. Mitt. 58, 50-53.
- Weimer, H. E. & Benjamin, D. C. (1965) Nature (London) 206. 1221-1222
- 61. Benjamin, D. C. & Weimer, H. E. (1966) Nature (London) 209, 1032-1033.
- 62. Andus, T., Gross, V., Tran-Thi, T.-A. & Heinrich, P. C. (1983) FEBS Lett. 151, 10-14.
- 63. Lin, T.-M. & Halbert, S. P. (1976) Science 193, 1249-1253.
- Lin, T.-M., Halbert, S. P. & Kiefer, D. (1976) J. Clin. Invest. 64. 57, 466-472.
- Stimson, W. H. & Blackstock, T. L. (1975) Experientia 31, 65. 371-373.
- Horne, C. H. W., Thomson, A. W., Towles, C. H., MacMil-66. lan, F. K. & Gibb, L. M. (1978) Scand. J. Immunol. 8, 75-80.
- Lundgren, E., Damber, M.-G., Roos, G., Von Schoultz, B., 67. Stigbrand, T., Nilsson, K. & Alexander, T. T. (1979) Int. J. Cancer 24, 45-48.
- Stimson, W. H., Farquharson, D. M., Shepherd, A. & Ander-68. son, J. M. (1979) J. Clin. Lab. Immunol. 2, 235-238.
- 69. Von Schoultz, B. & Stigbrand, T. (1982) in Pregnancy Proteins, eds. Grudzinskas, J. G., Teisner, B. & Seppälä, M. (Academic, Sydney), pp. 167-175.