Tümmler, B., Herrmann, U., Maass, G., & Eibl, H. (1984) *Biochemistry 23*, 4068–4074.
Wilkinson, D. A., & Nagle, J. F. (1984) *Biochemistry 23*, 1538–1541.
Wilkinson, D. A., & McIntosh, T. J. (1986) *Biochemistry 25*, 295–298.
Zaccai, G., Büldt, G., Seelig, A., & Seelig, J. (1979) *J. Mol. Biol. 134*, 693–706.

Biosynthesis and Molecular Cloning of Sulfated Glycoprotein 2 Secreted by Rat Sertoli Cells†

Michael W. Collard and Michael D. Griswold*
Biochemistry/Biophysics Program, Washington State University, Pullman, Washington 99164-4660
Received November 21, 1986; Revised Manuscript Received February 6, 1987

ABSTRACT: Sulfated glycoprotein 2 (SGP-2) is the major protein secreted by rat Sertoli cells. Pulse–chase labeling shows that SGP-2 is synthesized as a cotranslationally glycosylated 64-kDa precursor that is modified to a negatively charged 73-kDa form before intracellular cleavage to the mature 47- and 34-kDa subunits. A plasmid cDNA library was constructed from immunopurified mRNA, and a recombinant clone containing the entire protein coding sequence of SGP-2 was isolated. The 1857-nucleotide cDNA consists of a 297-nucleotide 5′ noncoding segment, a 1341-nucleotide coding segment, and a 219-nucleotide 3′ noncoding sequence. The 5′ noncoding region contains five ATG codons followed by four short open reading frames. The derived SGP-2 sequence has a molecular weight of 51 379 and contains six potential N-glycosylation sites. Proteolytic processing sites for the preproprotein were determined by amino-terminal sequencing of the isolated SGP-2 subunits. Northern blots show a wide tissue distribution for the 2.0-kb SGP-2 message, and computer sequence analysis indicates a significant relationship between SGP-2 and human apolipoprotein A-I.

Sertoli cells are the somatic component of the seminiferous epithelium. Because of their close association with developing germinal cells and their secretory nature, Sertoli cells are thought to provide both physical and biochemical support to the process of spermatogenesis. The formation of tight-junctional complexes between adjacent Sertoli cells creates an effective "blood–testis" barrier and results in a tubular fluid composition that is defined primarily by Sertoli cell secretion products (Waites, 1977).

Analysis of proteins from cultured cells has shown that the majority of secreted protein synthesis is directed toward the production of a single sulfated glycoprotein referred to as SGP-2[1] (previously DAG protein; Kissinger et al., 1982; Sylvester et al., 1984). SGP-2 is a disulfide-linked heterodimer whose reduced subunits migrate with mobilities of 47 and 34 kDa in SDS–polyacrylamide gels. SGP-2 contains 23.7% carbohydrate, and the N-linked oligosaccharides are extensively sulfated (Griswold et al., 1986). After secretion, SGP-2 can be detected by immunofluorescence on the acrosome and distal tail portion of mature spermatozoa (Sylvester et al., 1984). Although the precise function of SGP-2 has not been defined, the binding of the protein to spermatozoa and the extent to which it is produced in Sertoli cells and epididymis suggest that SGP-2 may play a critical role in spermatogenesis.

In this paper, we present evidence that Sertoli cells synthesize SGP-2 as a single preproprotein that is then post-translationally modified and cleaved before secretion to the extracellular space. We present the nucleotide sequence of a cDNA that encodes the entire amino acid sequence of the SGP-2 precursor and examine the synthesis of SGP-2 mRNA in testis and other tissues.

Materials and Methods

Cell Culture and Labeling Conditions. Sertoli cell cultures were prepared from 20-day-old rats as previously described (Dorrington & Fritz, 1975; Kissinger et al., 1982). Cells were plated onto 60-mm Falcon dishes and were maintained at 32 °C in Ham's F-12 medium supplemented with 0.1 mM dibutyryl-cAMP. The medium was changed after 2 days, and the cells were labeled after the fourth day of culture. Cells were pulsed-labeled with 200 μCi of [^{35}S]methionine in 0.5 mM of F-12 medium that lacked unlabeled methionine. At the end of each chase period, cells were lysed in 1.5 mL of buffer A (0.5% Triton X-100, 0.5% sodium deoxycholate, 150 mM NaCl, 50 mM Tris, pH 7.5, 5 mM EDTA, 0.5 mM phenylmethanesulfonyl fluoride) and frozen at −20 °C. The medium from the 1- and 4-h chase cells was centrifuged at 2000*g* to remove cellular debris and was frozen at −20 °C. Long-term labeling of Sertoli cell secreted proteins was achieved by incubating Sertoli cell cultures with 5 mL of

† This research was supported by National Institutes of Health Grant HD-10808 and by a grant (to M.W.C.) awarded by the Graduate School, Washington State University.
* Address correspondence to this author.

[1] Abbreviations: SGP-1, sulfated glycoprotein 1; SGP-2, sulfated glycoprotein 2; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; cAMP, adenosine cyclic 3′,5′-phosphate; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside; IPTG, isopropyl β-D-thiogalactopyranoside; PTH, phenylthiohydantoin; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; NBRF, National Biomedical Research Foundation; PIR, Protein Identification Resource; kDa, kilodalton(s).

0006-2960/87/0426-3297$01.50/0

Ham's F-12 medium that lacked unlabeled methionine and to which 500 μCi of [^{35}S]methionine has been added. The medium was then desalted as previously described (Sylvester et al., 1984).

Immunoprecipitation. Prior to immunoprecipitation, samples were thawed and 0.2-mL aliquots of cell lysates or medium were centrifuged at 8000*g* for 5 min to remove insoluble material. Each sample was mixed with 0.2 mL of buffer A and 3 μL of rabbit anti-SGP-2 serum and then incubated at room temperature for 1 h (Sylvester et al., 1984). Immune complexes were precipitated by the addition of 50 μL of a 10% w/v suspension of *Staphylococcus aureus* (Pansorbin, Calbiochem-Behring) for 30 min. The mixture was then centrifuged at 6000*g* for 3 min, and the pellet was washed 4 times by resuspension and centrifugation in buffer A. The immune complexes were solubilized by resuspension of the pellets in the appropriate electrophoresis sample buffer at 95 °C for 4 min, followed by a final centrifugation at 12000*g* for 4 min to remove the insoluble Pansorbin.

cDNA Synthesis and Selection of Recombinants. Immunopurified mRNA was used as a template for cDNA synthesis by the method of Land et al. (1981). The cDNA was cloned into the *Pst*I site of pUC13 by G-C tailing and then transfected into *Escherichia coli* JM105 (Viera & Messing, 1982). Plasmid DNA was purified from white colonies selected on X-gal/IPTG indicator plates and was sized by agarose gel electrophoresis. Plasmids containing large (>1-kb) inserts were immobilized on nitrocellulose and were used to hybrid select mRNA from total Sertoli cell mRNA (Parnes et al., 1981; Gurney et al., 1982).

DNA Sequence Analysis and SP6 Plasmid Construction. Appropriate restriction fragments from the plasmid designated p60 were subcloned into pTZ18U and pTZ19U plasmid vectors (United States Biochemicals), and single-stranded DNA templates were generated as outlined by the manufacturer. Some of the termini used for sequencing were generated by unidirectional digestion with 3′-exonuclease (Henikoff, 1984). The templates were sequenced by the dideoxy chain termination method as modified for use with reverse transcriptase (Sanger et al., 1977; Duncan, 1985). The *Hin*dIII/*Eco*RI restriction fragment from p60 was subcloned into the *Hin*dIII/*Eco*RI restriction site of the pSP64 plasmid (Promega Biotech). This construction resulted in the placement of the SP6 polymerase promoter 5′ to the cDNA sequence shown in Figure 4. RNA was synthesized from purified plasmid with SP6 polymerase according to protocols supplied by the manufacturer (Promega Biotech).

Northern Blot Analysis and RNA Isolation. RNA was isolated from cell cultures or whole tissue by the guanidine isothiocynate/LiCl procedure (Cathala et al., 1983). Poly(A+) RNA was obtained by oligo(dT)–cellulose chromatography and was fractionated and blotted as previously described (Aviv & Leder, 1972; Huggenvik et al., 1986). The blots were hybridized with 200 ng of ^{32}P-labeled nick translated cDNA probe (10^8 cpm/μg of plasmid p60). The removal of polyadenylation from Sertoli cell poly(A+) RNA utilized RNase H according to the protocol of Kleene et al. (1984).

In Vitro Translation and Endoglycosidase H Digestion. Approximately 1 μg of poly(A+) RNA or SP6 polymerase derived RNA was translated in a wheat germ extract system (Bethesda Research Laboratories) supplemented with 50 μCi of [^{35}S]methionine (New England Nuclear) per 30-μL reaction. The translation products from 3-μL aliquots were immunoprecipitated or analyzed directly by SDS–PAGE. For digestion with endoglycosidase H, immunoprecipitates were eluted from the Pansorbin pellet in 50 μL of 100 mM citrate, pH 5.8, 10 mM DTT, and 0.2% SDS. The supernatant was divided into equal portions, and 45 ng of endoglycosidase H (New England Nuclear) was added to one of the samples. Both samples were incubated 24 h at 37 °C and were then mixed with an equal volume of 2X electrophoresis sample buffer.

Immunopurification of SGP-2 mRNA. Polysomes were isolated from homogenates of adult rat testes by the Mg^{2+} precipitation method of Palmiter (1974). Antibodies to SGP-2 were purified, and polysomes were immunoadsorbed to protein A–Sepharose (Sigma) as described by Kraus and Rosenberg (1982). The eluted polysomes were then processed for poly(A+) RNA as previously described (Huggenvik et al., 1986).

Polyacrylamide Gel Electrophoresis. Proteins were analyzed by one-dimensional SDS–PAGE on slab gels containing 10% acrylamide (Laemmli, 1970). Two-dimensional electrophoresis was done by the method of O'Farrell (1975) with modifications previously described (Kissinger et al., 1982).

Protein Purification and Sequencing. SGP-2 was purified by two consecutive reverse-phase HPLC runs as previously described (Griswold et al., 1986). SGP-2 was reduced with 10 mM dithiothreitol and rechromatographed by HPLC under the previously described conditions. Three fractions were obtained which contained the large, unreduced, and small subunits of SGP-2. The fractions containing the large and small subunits were dialyzed against H_2O and lyophilized. The samples were then sequenced on an Applied Biosystems 470A protein sequencer according to the protocols supplied by the manufacturer.

Computer Sequence Analysis. The NBRF/PIR computer programs were run on a VAX 11/785 computer (Digital Equipment Corp.). The NBRF/PIR protein sequence data base (version 7.0), the GenBank data base (release 44; Bolt, Beranek and Newman, Inc.) and the EMBL data base (version 6.0) were used in sequence comparisons.

RESULTS

Biosynthesis of Subunits. Primary cultures of rat Sertoli cells were grown in media supplemented with dibutyryl-cAMP and were labeled with [^{35}S]methionine on the fourth day of culture. Proteins isolated from the cultures were fractionated by SDS–PAGE under reducing conditions and were subsequently fluorographed. Lane 1 of Figure 1 shows the profile of proteins secreted into culture media after a 24-h labeling period. The majority of the [^{35}S]methionine labeled protein is represented by the reduced 47- and 34-kDa subunits of SGP-2. The two other prominent Sertoli cell secreted proteins visible in lane 1 are SGP-1 (70 kDa) and testicular transferrin (76 kDa) (Kissinger et al., 1982). A polyclonal antiserum to SGP-2 (Sylvester et al., 1984) specifically immunoprecipitates the mature protein from total secreted proteins (Figure 1, lane 2) and was used in the isolation of SGP-2 precursors in pulse–chase experiments.

Sertoli cell cultures were pulse-labeled with [^{35}S]methionine for 15 min, after which the medium was removed. Cells were then immediately lysed or incubated an additional 1–4 h with medium containing an excess of nonradioactive methionine. At the end of each chase period, both cells and medium were collected for immunoprecipitation with SGP-2 antiserum. A single 64-kDa protein was immunoprecipitated from Sertoli cells that had been incubated with [^{35}S]methionine for 15 min (Figure 1, lane 3). After a 1-h chase, the precursor decreased in intensity while three other immunoprecipitable bands became apparent (Figure 1, lane 4). A slight shift in precursor molecular weight from 64 000 to 62 000 was also observed,

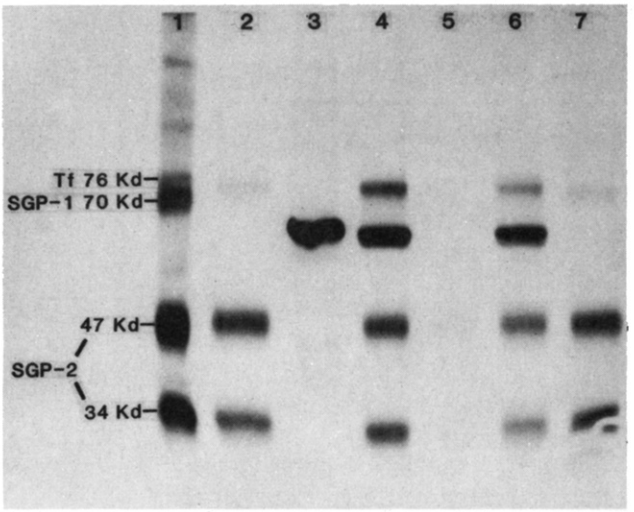


FIGURE 1: Pulse–chase analysis of SGP-2 synthesis in cultured rat Sertoli cells. The figure is a fluorograph of [^{35}S]methionine-labeled proteins separated on a SDS–10% polyacrylamide gel under reducing conditions. Lane 1 depicts the proteins secreted into culture medium after a 24-h labeling period. Lanes 2–7 are immunoprecipitations of culture medium or cell lysates using anti-SGP-2 serum. Lane 2, culture media after a 24-h labeling period. Lane 3, cell lysate after a 15-min pulse-labeling period. Cell lysate (lane 4) and culture medium (lane 5) after a 15-min pulse-labeling period followed by a 1-h chase with unlabeled methionine. Cell lysate (lane 6) and culture medium (lane 7) after a 15-min pulse-labeling period followed by a 4-h chase with unlabeled methionine.

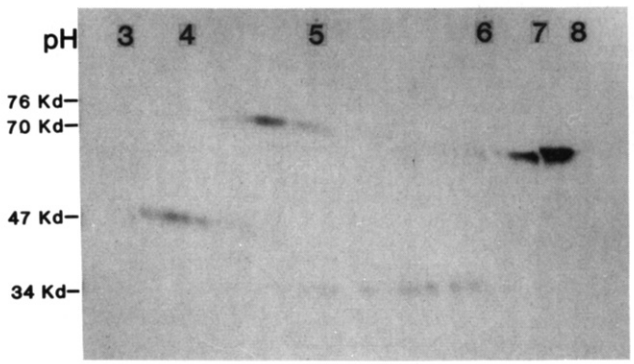


FIGURE 2: Fluorograph of the two-dimensional electrophoretic separation of proteins shown in lane 4 of Figure 1. Sertoli cells were pulse-labeled for 15 min with [^{35}S]methionine and then chased for 1 h in the presence of unlabeled methionine. Cell lysates were then immunoprecipitated with anti-SGP-2 serum.

which possibly indicates cleavage of the leader peptide by this time. Only a trace amount of radiolabeled SGP-2 was detected in the culture medium after 1 h of chase (Figure 1, lane 5).

Of the three additional protein bands observed in cells after 1 h of chase, two of the bands comigrate with the mature 47- and 34-kDa reduced subunits of secreted SGP-2. The third band migrates with a mobility of 73 kDa. After a 4-h chase, the majority of the mature SGP-2 (47- and 34-kDa subunits) had been secreted into the culture medium (Figure 1, lane 7). The persistence of the 62-kDa cellular precursor 4 h after synthesis (Figure 1, lane 6) is possibly due to an incomplete chase of the [^{35}S]methionine. Attempts to use cycloheximide to increase the efficiency of the chase resulted in a rapid degradation of newly synthesized SGP-2 protein (data not shown). The immunoprecipitable 73-kDa precursor protein also persisted after 4 h of chase and, to a small extent, was secreted into the medium.

When SGP-2 proteins from cells that were chased for 1 h (Figure 1, lane 4) were analyzed by two-dimensional gel

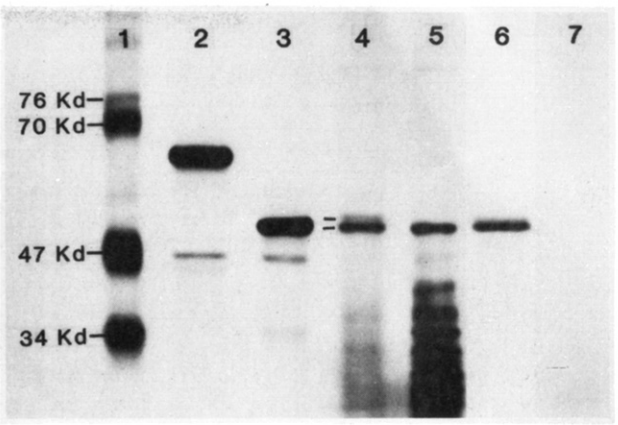
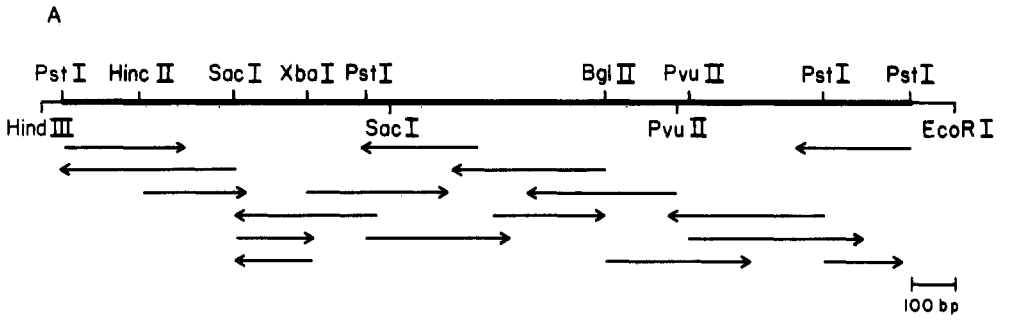


FIGURE 3: Fluorograph of [^{35}S]methionine-labeled proteins treated with endoglycosidase H or synthesized by in vitro translation. Proteins were electrophoresed on a SDS–10% polyacrylamide gel under reducing conditions. Lane 1, Sertoli cell secreted proteins (as markers). Lanes 2 and 3, SGP-2 proteins immunoprecipitated from Sertoli cells that had been pulse-labeled for 15 min with [^{35}S]methionine (as described for lane 3 of Figure 1). The immunoprecipitated proteins were divided into equal portions and were incubated in the absence (lane 2) or presence (lane 3) of endoglycosidase H for 24 h at 37 °C. Lanes 4–7 are [^{35}S]methionine-labeled in vitro translation products obtained after the addition of the following RNAs: Lane 4, immunopurified SGP-2 mRNA. Lane 5, SP6 polymerase derived RNA obtained from a plasmid containing the SGP-2 cDNA sequence. Lane 6, RNA selected from total Sertoli cell poly(A+) RNA by hybridization to immobilized p60 plasmid. The translation products were then immunoprecipitated with anti-SGP-2 serum. Lane 7, translation containing no added RNA. The two bars between lanes 3 and 4 indicate the major 50-kDa and minor 51-kDa translation products. Some compression of these two proteins occurs in lane 6 due to the presence of unlabeled immunoglobulin G.

electrophoresis (Figure 2), it was found that the 62-kDa cellular precursor possessed a p*I* of approximately 7.5, while the cellular 47- and 34-kDa proteins displayed acidic p*I*'s and charge heterogeneties similar to those previously observed for secreted SGP-2 protein (Kissinger et al., 1982). In addition, the 73-kDa band displayed an average p*I* of 4.6 that was intermediate to those of the 47- and 34-dKa subunits and also showed similar charge heterogeneity. It should be noted that the 73-kDa band was distinct in both size and immunoreactivity from the 70-kDa SGP-1 protein (Figure 1 and unpublished observations).

Construction and Sequence of SGP-2 cDNA. Polysomes were prepared from adult rat testes, and RNA enriched for SGP-2 transcripts was obtained by immunoadsorption to protein A–Sepharose. In vitro translation of the SGP-2-enriched poly(A+) RNA resulted in the production of a major 50-kDa protein and also a minor protein species that migrated at approximately 51 kDa (Figure 3, lane 4). Incubation of the Sertoli cell pulse-labeled 64-kDa precursor (Figure 3, lane 2; Figure 1, lane 3) with endoglycosidase H also produced a 50-kDa protein (Figure 3, lane 3), suggesting that the 50-kDa protein observed in the in vitro translations indicated an enrichment for SGP-2 mRNA sequences.

Double-stranded cDNA was synthesized from SGP-2 enriched mRNA by the method of Land et al. (1981). The tailed cDNA was annealed into the *Pst*I site of pUC13 and was used to transform the JM105 strain of *E. coli*. A positive clone designated p60 was initially identified by hybrid-select translation of SGP-2 mRNA from total Sertoli cell mRNA (Figure 3, lane 6) and was later confirmed by DNA and protein sequencing (see below). Subcloning of the p60 insert into a plasmid containing an SP6 RNA polymerase promoter allowed in vitro production of SGP-2 mRNA. In vitro translation of the SP6 polymerase derived RNA resulted in



B

```
-297  CTC AAG CTC ACA CTG GCT GGA CTT CCT CGC CAT GAC AGT CTG TAC CTC TAA CTG ATC CCA GGG [ATG ATA CCA CCT ACA TTT GGG   -214
-213  GTG GTT CTT CTC GCC TCA GTT AAA CCT CTC TGG GAG CAC CAT CAC AGA CAC CCA CAG AAG TTT GTT CCC TAG][ATG ATT CTA GGT   -130
-129  CCT GTG GAG TTG ACA AGA TTG ACC ATC ACG CTC TCA GCA ATC GGG TGA] AGT AAA CAC CAC CGT TGT CTC CAT GGA AAT GCT TAA   -46

                                                                  Met Lys Ile Leu Leu Leu Cys Val Ala Leu Leu Leu Thr  13
-45   CTA CGG CTT GCT AGT AAG GAC TCC AGA CTC CAA AGA GGC CAC ACC ATG AAG ATT CTC CTG CTG TGT GTG GCA CTG CTG CTG ACC   39

 14   Trp Asp Asn Gly Met Val Leu Gly↓Glu Gln Glu Phe Ser Asp Asn Glu Leu Gln Glu Leu Ser Thr Gln Gly Ser Arg Tyr Val  41
 40   TGG GAC AAT GGC ATG GTC CTG GGA GAG CAG GAG TTC TCT GAC AAT GAG CTC CAA GAA CTG TCC ACT CAA GGA AGT AGG TAT GTT  123

 42   Asn Lys Glu Ile Gln Asn Ala Val Gln Gly Val Lys His Ile Lys Thr Leu Ile Glu Lys Thr Asn Ala Glu Arg Lys Ser Leu  69
124   AAT AAG GAG ATT CAG AAC GCC GTC CAG GGG GTG AAG CAC ATA AAG ACC CTC ATA GAA AAA ACC AAC GCA GAG CGC AAG TCC CTG  207

 70   Leu Asn Ser Leu Glu Glu Ala Lys Lys Lys Lys Glu Gly Ala Leu Asp Asp Thr Arg Asp Ser Glu Met Lys Leu Lys Ala Phe  97
208   CTC AAC AGT TTA GAG GAA GCC AAA AAG AAG AAA GAG GGT GCT CTA GAT GAC ACC AGG GAT TCT GAA ATG AAG CTG AAG GCT TTC  291

 98   Pro Glu Val Cys Asn Glu Thr Met Met Ala Leu Trp Glu Glu Cys Lys Pro Cys Leu Lys His Thr Cys Met Lys Phe Tyr Ala  125
292   CCG GAA GTG TGT AAC GAG ACC ATG ATG GCC CTC TGG GAA GAG TGT AAG CCC TGC CTG AAG CAC ACC TGC ATG AAG TTC TAC GCA  375

126   Arg Val Cys Arg Ser Gly Ser Gly Leu Val Gly Arg Gln Leu Glu Glu Phe Leu Asn Gln Ser Ser Pro Phe Tyr Phe Trp Met  153
376   CGC GTC TGC AGG AGC GGC TCG GGG CTG GTT GGT CGC CAG CTA GAG GAG TTT CTG AAC CAG AGC TCA CCC TTC TAC TTC TGG ATG  459

154   Asn Gly Asp Arg Ile Asp Ser Leu Leu Glu Ser Asp Arg Gln Gln Ser Gln Val Leu Asp Ala Met Gln Asp Ser Phe Thr Arg  181
460   AAC GGG GAC CGC ATC GAC TCC CTG CTG GAG AGT GAC CGG CAG CAG AGC CAA GTC CTA GAT GCT ATG CAG GAC AGC TTC ACT CGG  543

182   Ala Ser Gly Ile Ile His Thr Leu Phe Gln Asp Arg Phe Phe Thr His Glu Pro Gln Asp Ile His His Phe Ser Pro Met Gly  209
544   GCG TCT GGC ATC ATA CAT ACG CTT TTC CAG GAC CGG TTC TTC ACC CAT GAG CCC CAG GAC ATC CAC CAT TTC TCC CCC ATG GGC  627

210   Phe Pro His Lys Arg Pro His Phe Leu Tyr Pro Lys Ser Arg Leu Val Arg↓Ser Leu Met Pro Leu Ser His Tyr Gly Pro Leu  237
628   TTC CCA CAC AAG CGG CCT CAT TTC TTG TAC CCC AAG TCC CGC TTG GTC CGC AGC CTC ATG CCT CTC TCC CAC TAC GGG CCT CTG  711

238   Ser Phe His Asn Met Phe Gln Pro Phe Phe Asp Met Ile His Gln Ala Gln Gln Ala Met Asp Val Gln Leu His Ser Pro Ala  265
712   AGC TTC CAC AAC ATG TTC CAG CCT TTC TTT GAT ATG ATA CAC CAG GCT CAA CAG GCC ATG GAC GTC CAG CTC CAT AGC CCA GCT  795

266   Leu Gln Phe Pro Asp Val Asp Phe Leu Lys Glu Gly Glu Asp Asp Pro Thr Val Cys Lys Glu Ile Arg His Asn Ser Thr Gly  293
796   TTA CAG TTC CCG GAT GTG GAT TTC TTA AAA GAA GGT GAA GAT GAC CCG ACA GTG TGC AAG GAG ATC CGC CAT AAC TCC ACA GGA  879

294   Cys Leu Lys Met Lys Gly Gln Cys Glu Lys Cys Gln Glu Ile Leu Ser Val Asp Cys Ser Thr Asn Asn Pro Ala Gln Ala Asn  321
880   TGC CTG AAG ATG AAG GGC CAG TGT GAG AAG TGC CAA GAG ATC TTG TCT GTG GAC TGT TCG ACC AAC AAT CCT GCC CAG GCT AAC  963

322   Leu Arg Gln Glu Leu Asn Asp Ser Leu Gln Val Ala Glu Arg Leu Thr Gln Gln Tyr Asn Glu Leu Leu His Ser Leu Gln Ser  349
964   CTG CGC CAG GAG CTA AAC GAC TCG CTC CAG GTG GCT GAG AGG CTG ACC CAG CAG TAC AAC GAG CTG CTT CAT TCC CTC CAG TCC  1047

350   Lys Met Leu Asn Thr Ser Ser Leu Leu Glu Gln Leu Asn Asp Gln Phe Thr Trp Val Ser Gln Leu Ala Asn Leu Thr Gln Gly  377
1048  AAG ATG CTC AAC ACC TCA TCC CTG CTG GAA CAG CTG AAC GAC CAG TTC ACG TGG GTG TCC CAG CTG GCT AAC CTC ACA CAG GGC  1131

378   Asp Asp Gln Tyr Leu Arg Val Ser Thr Val Thr Thr His Ser Ser Asp Ser Glu Val Pro Ser Arg Val Thr Glu Val Val Val  405
1132  GAT GAC CAG TAC CTT CGG GTC TCC ACA GTG ACA ACC CAT TCT TCT GAC TCA GAA GTC CCC TCT CGT GTC ACT GAG GTG GTG GTG  1215

406   Lys Leu Phe Asp Ser Asp Pro Ile Thr Val Val Leu Pro Glu Glu Val Ser Lys Asp Asn Pro Lys Phe Met Asp Thr Val Ala  433
1216  AAG CTG TTT GAC TCT GAC CCC ATC ACA GTG GTG TTA CCA GAA GAA GTC TCC AAG GAT AAC CCT AAG TTT ATG GAC ACA GTG GCA  1299

434   Glu Lys Ala Leu Gln Glu Tyr Arg Arg Lys Ser Arg Met Glu End
1300  GAG AAA GCG CTA CAG GAA TAC CGC AGG AAA AGC CGC ATG GAA TGA GAC AGA AGC ATC AGT TTT CTA TAT GTA GGA GTC TCA AGG  1383

3844  AGG GAA TCT CCC AGC TTT CCG AGG TTG CTG CAG ACC CCT AGA GAA CTC ACA TGT CTC CAG CGC CTA GGC CTC CAC CCC AGC AGC  1467

1468  CTC TCC TTC CTC TGG GTT CTG TAC TCT AAT GCC TGC ACT TGA TGC TCT GGG AAG AAC TGC TTC CCC CAC GCA ACT AAT CCA ATA  1551

1552  AAG CAC CTT   1560
```

FIGURE 4: Sequencing strategy and complete nucleotide sequence of rat SGP-2 cDNA. (A) Nucleotide sequence determined by the dideoxy chain termination method. The arrows represent the strand and extent of the sequence determinations of each subcloned DNA fragment. Two of the termini used for sequencing were generated by unidirectional 3′-exonuclease digestion. The *Eco*RI and *Hin*dIII restriction sites are included to orient the cDNA fragment in the pUC13 multiple cloning site. (B) Nucleotide and derived protein sequence of the p60 cDNA insert. Nucleotide and amino acid numbering occurs in the margins. Amino acid residues confirmed by N-terminal protein sequencing are indicated by overlines. Arrows indicate sites of proteolytic processing. Closed circles indicate potential N-glycosylation sites, and open circles indicate cysteine residues. A polyadenylation consensus sequence (AATAAA) is underlined. Four small open reading frames occuring within the 5′ noncoding region are indicated by brackets. Not shown are 14 G residues at the 5′ end and 13 A residues followed by 16 C residues at the 3′ end.

the production of a 50-kDa protein (Figure 3, lane 5) but did not produce the minor 51-kDa protein that was observed in translations of immunopurified SGP-2 mRNA or hybrid-selected SGP-2 mRNA (Figure 3, lanes 4 and 6). The smaller proteins observed in the SP6 translation are probably due to the presence of incomplete RNA transcripts.

The strategy used to sequence the 1857-nucleotide insert in p60 is illustrated in Figure 4A, and the complete nucleotide sequence is shown in Figure 4B. An opening reading frame, designated to begin at nucleotide 1, continues through nucleotide 1341 and terminates with a single TGA codon. There are 297 nucleotides of 5′ noncoding sequence and 219 nucleotides of 3′ noncoding sequence, with a polyadenylation consensus sequence at position 1548. Five ATG codons occur 5′ to the ATG codon assigned the number 1 position, and these occur at positions −266, −234, −241, −59, and −53, respectively. In-frame termination codons occur 7, 30, 19, 20, and 18 codons downstream from each of these ATG sites. Several criteria were used in our assignment of the codon that is used for initiation of SGP-2 protein synthesis. The ATG codon at position 1 initiates the only open reading frame that could produce a protein similar in length to that observed for the in vitro translation product. The derived molecular weight of the 447 amino acid long protein is 51 379. Nucleotides −3 to +4 conform closely to the translation initiation consensus sequence (ACCATGG) described by Kozak (1986). The first 21 amino acids of the derived sequence have a composition typical of that expected for a leader peptide sequence. This

includes a positively charged amino acid directly following the initiator methionine, a hydrophobic midsection, and a small amino acid at the putative position of cleavage by leader peptidase (von Heijne, 1985).

Amino-terminal protein sequencing of mature SGP-2 was done in order to determine the sites of proteolytic processing and to also confirm the identification of the SGP-2 cDNA clone. SGP-2 was purified as previously described (Griswold et al., 1983), and the small and large subunits were isolated as described under Materials and Methods. Approximately 100 pmol of the small subunit and 50 pmol of the large subunit were subjected to gas-phase sequencing on an Applied Biosystems 470A protein sequencer. Positions 22–45 and 224–241 of the derived protein sequence correspond exactly with the first 24 and 18 PTH-amino acid derivatives obtained from Edman degradation of the small and large subunits, respectively (Figure 3). These results confirm Gly-21 as the site of leader peptide cleavage and also indicate that conversion of pro-SGP-2 to the mature protein occurs by protoelytic cleavage at Arg-226. Subunits produced by these cleavages have derived protein molecular weights of 23 921 and 25 176. Of the six potential N-glycosylation sites (Asn-X-Ser and Asn-X-Thr) that occur in the derived sequence, two are found in the small subunit while four are found in the large subunit. There are 11 cysteine residues in the derived protein sequence, one of which is contained in the leader peptide. The remaining 10 residues are equally divided between the two subunits and appear distributed near potential sites of glycosylation (Figure 4).

Northern Blot Analysis. Expression of SGP-2 mRNA in a variety of tissues was examined by Northern blot analysis. Poly(A+) RNA was isolated from testis, epididymis, liver, brain, kidney, and spleen of adult male rats. Poly(A+) RNA was also isolated from the mammary glands of lactating female rats and from Sertoli cell cultures of 20-day-old animals. Three micrograms of mRNA from the various sources were fractionated on a denaturing 1.5% agarose gel and subsequently transferred to nitrocellulose. The blot was then hybridized with nick translated p60 plasmid. Figure 5A shows an autoradiogram of the nitrocellulose blot and indicates the relative levels of SGP-2 mRNA expression among the tissues. The highest levels of SGP-2 mRNA occur in epididymis and cultured Sertoli cells. Densitometric scanning of the autoradiogram indicated that testis and brain produced 96% and 67%, respectively, the level of mRNA observed for liver. Longer exposure of the autoradiogram showed detectable levels of SGP-2 mRNA in kidney, spleen, and mammary gland.

Expression of testicular SGP-2 mRNA as a function of age was also examined by Northern blot analysis (Figure 5B). Poly(A+) RNA was isolated from the testes of rats varying in age from 5 to 72 days and was analyzed as described above. Detectable levels of SGP-2 mRNA were observed in 5-day-old rats, and a steady increase in message occurred up to 20 days of age. There was an apparent decline (see Discussion) in mRNA levels between 20 and 42 days of age, while a constant level of mRNA was observed between 42 and 72 days of age.

Northern analysis was used to measure the molecular size of SGP-2 mRNA before and after the removal of polyadenylation. Sertoli cell poly(A+) RNA was hybridized to oligo(dT) and was then incubated in the absence or presence of RNase H. The samples were then analyzed by Northern blot as described above. The molecular size of polyadenylated SGP-2 mRNA was estimated to be 2000 nucleotides, while the RNase H treated sample displayed a molecular size of approximately 1840 nucleotides (Figure 5C). This size was in good agreement with the 1857-nucleotide-long cDNA.

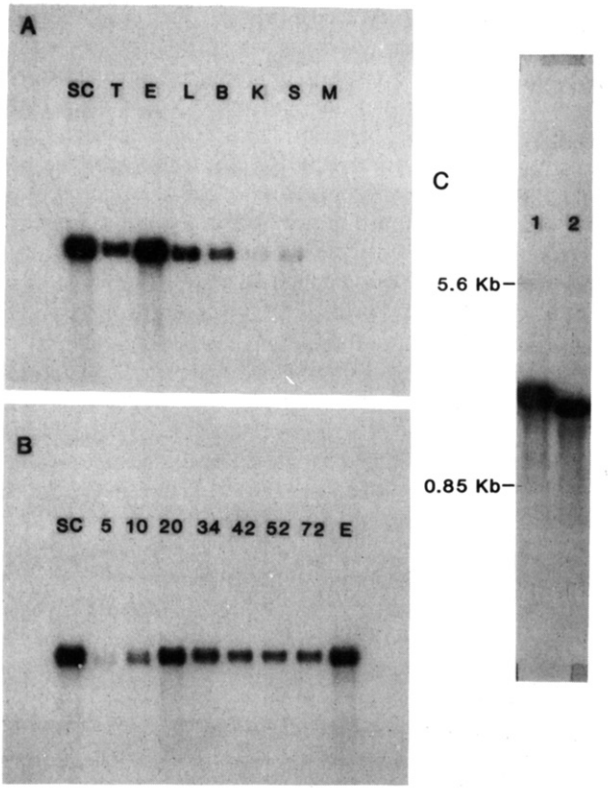


FIGURE 5: Northern blot analysis of SGP-2 mRNA. Three micrograms of poly(A+) RNA was loaded in each lane, and the nitrocellulose blots were hybridized with nick translated p60 plasmid. (A) RNA isolated from rat Sertoli cell cultures (SC), testis (T), epididymis (E), Liver (L), brain (B), kidney (K), spleen (S), or mammary gland (M). (B) RNA isolated from rat Sertoli cell culture (SC), testis from 5-, 10-, 20-, 34-, 42-, 52-, and 72-day-old rats, or rat epididymis (E). (C) RNA isolated from rat Sertoli cell culture before (lane 1) and after (lane 2) removal of polyadenylation with oligo(dT) and RNase H.

Discussion

SGP-2 is the major protein secreted by cultured Sertoli cells. The SGP-2 subunit molecular weights of 47 000 and 34 000 reported here are the result of more recent measurements by using a variety of SDS–PAGE techniques and are larger than those previously described (Kissinger et al., 1982; Sylvester et al., 1984). We have used pulse–chase labeling techniques to show that SGP-2 is synthesized as a single glycosylated precursor with an SDS–PAGE mobility of 64 000 daltons. A 1-h chase of the 64-kDa precursor resulted in the production of three additional protein bands. Two of these bands comigrated with the 47- and 34-kDa subunits of mature SGP-2 while the third protein band migrated with a mobility of 73 kDa. Analysis of these proteins by two-dimensional gel electrophoresis showed that the precursor (62 kDa after removal of the signal sequence) was found as a single more basic species, while the 73-, 47-, and 34-kDa proteins appeared as acidic and multiply charged isomers (Figure 2). Previous work has shown that the acidic p*I* and charge heterogeneity of SGP-2 is due to sialation and sulfation of carbohydrate moieties (Griswold et al., 1986; Kissinger et al., 1982). Our data suggest that the 62-kDa precursor undergoes carbohydrate sialation and sulfation to form the acidic 73-kDa protein which is subsequently cleaved to the 47- and 34-kDa subunits of mature SGP-2. The large change in SDS–PAGE mobility upon modification of the 62-kDa precursor most likely reflects aberrant electrophoretic migration due to sulfated carbohy-

drate rather than a large change in mass. Our contention that carbohydrate modification occurs prior to proteolytic cleavage is supported by the observation that subunits smaller than 47 and 34 kDa (representing cleavage products from the 62-kDa precursor) do not occur in pulse-labeled SGP-2 protein (Figure 1). Proteolytic cleavage of SGP-2 does not appear to be a prerequisite for cellular secretion since a small amount of the pro-SGP-2 was secreted into culture medium (Figure 1). Pro-SGP-2 has also been detected in the rete testis fluid from intact animals (unpublished data), suggesting that its secretion may be a normal physiological process.

In vitro translation of immunopurified SGP-2 mRNA resulted in the production of an unglycosylated 50-kDa protein product. This molecular weight was in good agreement with that obtained by endoglycosidase H treatment of the 64-kDa SGP-2 precursor and also with the molecular weight obtained from the derived amino acid sequence. The estimate of 22% carbohydrate determined by SDS–PAGE mobilities for the glycosylated and unglycosylated precursor is similar to the 23.7% carbohydrate content found by chemical analysis of mature SGP-2 (Griswold et al., 1986). If the carbohydrate occurring on the SGP-2 64-kDa precursor is a typical $(Glc)_3(Man)_9(GlcNAc)_2$ core structure ($M_r \sim 2350$), then the minimum number of N-glycosylation sites required to produce the 14-kDa difference between unglycosylated and glycosylated SGP-2 is six. It therefore seems likely that all six potential N-glycosylation sites occurring in the SGP-2 sequence will be glycosylated. Since the derived protein molecular weights are quite similar for the two subunits, a differential glycosylation could explain the large size difference observed for the mature subunits on SDS–polyacrylamide gels.

In addition to a major 50-kDa protein, a minor protein species of 51 kDa was observed in translations of immunopurified mRNA and also in immunoprecipitations of translation products from hybrid-selected mRNA (Figure 3). The origin of this protein band is currently unknown. If the 51- and 50-kDa in vitro translation products were occurring due to occasional protein initiation at upstream ATG sites or perhaps primary initiation at the ATG site within the leader sequence (nucleotide +52), then RNA derived from the cloned cDNA sequence should also produce two in vitro translation products. However, only the 50-kDa protein species was observed in in vitro translations of SP6 polymerase derived mRNA. This result is consistent with the presence of in-frame stop codons downstream from each of the 5′ noncoding ATG sites and suggests the possibility of a minor species of mRNA coding for a larger isoform of SGP-2 protein. To date, only one species of SGP-2 mRNA has been detected in Northern blots of total testicular message and only one size of prepro-SGP-2 has been observed in methionine-labeled Sertoli cells. The relationship of the 51-kDa in vitro translation product to the normal cellular production of SGP-2 remains to be determined.

The overall structure of the 5′ noncoding region in rat SGP-2 mRNA is remarkably similar to those described for yeast GCN4 and murine ornithine decarboxylase mRNAs (Mueller & Hinnebusch, 1986; Khana & Nathans, 1985). Both GCN4 and ornithine decarboxylase contain four short open reading frames upstream from their protein initiation sites. The upstream AUG codons in GCN4 are responsible for the translational repression of GCN4 protein production, and the fourth AUG codon (5′ proximal) also appears necessary for efficient GCN4 expression under starvation conditions (Mueller & Hinnebusch, 1986). Selective inhibition of mRNA translation has also been described for the polyamine repression of ornithine decarboxylase synthesis (Holtta & Pohjanpelto, 1986). Comparison of the 5′ noncoding region of SGP-2 to those of GCN4 and ornithine decarboxylase shows several interesting differences. The 5′ open reading frames in SGP-2 tend to be much longer than those reported for GCN4 or ornithine decarboxylase. Also, the fourth open reading frame in SGP-2 overlaps with the SGP-2 coding sequence and terminates one nucleotide out of frame from the ATG codon used to initiate SGP-2 synthesis. Kozak (1986) has demonstrated that a purine in a position three nucleotides upstream from an ATG codon has a dominant effect on initiation by eukaryotic ribosomes. Three of the five ATG codons occurring in the 5′ noncoding region of SGP-2 display a G at the −3 position, including one of the two ATG codons that initiate the overlapping reading frame. This raises the possibility that small peptides up to 30 amino acids long may actually be synthesized from the 5′ noncoding region of SGP-2 mRNA. It will be of interest to determine the effects, if any, these upstream ATG sites have on SGP-2 synthesis in Sertoli cells.

Northern blot analysis showed that, after removal of polyadenylation, the molecular size of SGP-2 mRNA was approximately 1836 nucleotides. The 1858-nucleotide cDNA shown in Figure 4 therefore represents the majority, if not all, of the SGP-2 mRNA sequence. SGP-2 mRNA was found in a wide distribution of tissues (Figure 5A), with the highest levels occurring in epididymis and Sertoli cell culture. In situ hybridization data hs confirmed that testicular expression of SGP-2 mRNA is confined to Sertoli cells (Morales & Griswold, 1987). When the relative testicular levels of SGP-2 mRNA were examined from various aged animals, an apparent decrease occurred between 20 and 42 days of age (Figure 5B). Meiosis in the male rat first occurs at approximately 20 days of age. At this point, the number of Sertoli cells has reached a maximum, and rapid proliferation of germ cells begins and does not plateau until approximately 35 days of age. Since a constant amount of mRNA (3 μg) was loaded in each of the lanes in Figure 5, the apparent decline in SGP-2 mRNA production between days 20 and 42 more likely reflects the large increase in germinal cell mRNA occurring at this time.

The SGP-2 cDNA sequence was screened against release 7.0 of the Protein Identification Resource (PIR) data base in order to determine if SGP-2 shared sequence homology with any previously described proteins. The computer programs SEARCH and RELATE described by Dahoff et al. (1983) were used in this process. Sequences receiving the highest SEARCH scores were subjected to further analysis using RELATE. The RELATE program provides probability data on the homology of two proteins and can establish distant relationships that are not readily observed by direct comparison. The highest scoring sequence from the RELATE analysis was human apolipoprotein A-I. The rat prepro-SGP-2 sequence was compared to the human proapolipoprotein A-I precursor by using the mutation data scoring matrix (250 PAMs) and a fragment length of 12. The segment comparison score was calculated from 100 runs of the randomized sequences. A RELATE score of 4.1 standard deviation units was obtained, which indicated that the probability these sequences are related by chance is less than 2.1×10^{-5} (Dayhoff et al., 1983). This RELATE score is similar to those obtained by sequence comparison between other members of the apolipoprotein family (Boguski et al., 1986). The relationship of SGP-2 with the apolipoproteins is also consistent with several physical properties observed for SGP-2 protein. Despite its high degree of sulfation and glycosylation, SGP-2 elutes with a hydrophobic profile during HPLC iso-

lation and tends to aggregate upon purification (Griswold et al., 1986). Since the majority of known Sertoli cell secretion products are involved in transport, e.g., ceruloplasmin, transferrin, and androgen binding protein, it is tempting to speculate that SGP-2 could be involved in lipid transport.

ACKNOWLEDGMENTS

We thank Alice Karl for cell culture assistance and Jane Peterson for help in preparing the manuscript. We also thank Dr. Paul Bishop and James Tsuruta for their assistance in protein sequencing.

REFERENCES

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408.
Boguski, M. S., Elshourbagy, N. A., Taylor, J. M., & Jeffrey, I.G. (1986) *Methods Enzymol. 128*, 753.
Cathala, G., Savouret, J., Mendez, B., West, B. L., Karin, M., Martial, J. A., & Baxter, J. D. (1983) *DNA 2*, 329.
Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983) *Methods Enzymol. 91*, 524.
Dorrington, J. H., & Fritz, I. B. (1975) *Endocrinology (Philadelphia) 96*, 879.
Duncan, C. H. (1985) *N. Engl. Nucl. Prod. News 4(3)*, 6.
Griswold, M. D., Roberts, K. P., & Bishop, P. D. (1986) *Biochemistry 25*, 7265.
Gurney, T., Jr., Sorenson, D. K., Gurney, E. G., & Wills, N. M. (1982) *Anal. Biochem. 125*, 80.
Henikoff, S. (1984) *Gene 28*, 351.
Holtta, E., & Pohjanpelto, P. (1986) *J. Biol. Chem. 261*, 9502.
Huggenvik, J. I., Idzerda, R. L., Haywood, L., Lee, D. C., McKnight, G. S., & Griswold, M. D. (1987) *Endocrinology (Baltimore) 120*, 332.
Kahana, C., & Nathans, D. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1673.
Kissinger, C., Skinner, M. K., & Griswold, M. D. (1982) *Biol. Reprod. 27*, 233.
Kleene, K. C., Distel, R. J., & Hecht, N. B. (1984) *Dev. Biol. 105*, 71.
Kozak, M. (1986) *Cell (Cambridge, Mass.) 44*, 283.
Kraus, J. P., & Rosenberg, L. E. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4015.
Laemmli, U. K. (1970) *Nature (London) 227*, 680.
Land, H. M., Grez, H., Hansen, H., Lindermaier, W., & Schuetz, G. (1981) *Nucleic Acids Res. 9*, 2251.
Morales, C., & Griswold, M. D. (1987) *Biol. Reprod.* (in press).
Mueller, P. P., & Hinnebusch, A. G. (1986) *Cell (Cambridge, Mass.) 45*, 201.
O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007.
Palmiter, R. D. (1974) *Biochemistry 13*, 3606.
Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E., & Seidman, J. G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2253.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463.
Sylvester, S. R., Skinner, M. K., & Griswold, M. D. (1984) *Biol. Reprod. 31*, 1087.
Vieira, J., & Messing, J. (1982) *Gene 19*, 259.
von Heijne, G. (1985) *J. Mol. Biol. 184*, 99.

Interaction of *trans*-Diamminedichloroplatinum(II) with DNA: Formation of Monofunctional Adducts and Their Reaction with Glutathione†

Alan Eastman* and Michael A. Barry
Eppley Institute for Research in Cancer, University of Nebraska Medical Center, Omaha, Nebraska 68105
Received October 29, 1986

ABSTRACT: Bifunctional reactions with DNA are responsible for the toxic action of the cancer chemotherapeutic drug *cis*-diamminedichloroplatinum(II) (*cis*-DDP). Thiourea has previously been used to trap transient monofunctional adducts in DNA before they rearrange to the toxic lesions. In these studies, thiourea was used to quantify the monofunctional adducts produced by the ineffective isomer *trans*-DDP. Rather than trapping monofunctional adducts, thiourea labilized them from DNA. At short time periods, 85% of *trans*-DDP bound to double-stranded DNA as monofunctional adducts of deoxyguanosine. Rearrangement to bifunctional adducts in double-stranded DNA was 50% complete in 24 h but was much more rapid in single-stranded DNA with 100% complete rearrangement in 24 h. The ineffectiveness of *trans*-DDP therefore results from a high proportion of monofunctional adducts in DNA that rearrange very slowly to toxic bifunctional adducts. The persistent monofunctional adducts react rapidly with glutathione, which would further reduce their potential toxicity by preventing them from rearranging to more toxic bifunctional adducts.

The cancer chemotherapeutic agent *cis*-diamminedichloroplatinum(II) (*cis*-DDP)[1] is firmly established as a drug of high potency in the treatment of a variety of tumors. Reaction with DNA appears essential to the toxic action of the drug (Roberts & Thompson, 1979). A complex variety of lesions occur in DNA, including monofunctional adducts, interstrand cross-links, intrastrand cross-links, DNA–protein cross-links and DNA–glutathione cross-links [reviewed in Eastman (1987a)]. The bifunctional reactions are critical to the toxic action as

† Supported by National Cancer Institute Research Grants CA36039 and CA00906 and Cancer Center Support Grant CA36727.
* Address correspondence to this author.

[1] Abbreviations: *cis*-DDP, *cis*-diamminedichloroplatinum(II); *trans*-DDP, *trans*-diamminedichloroplatinum(II); HPLC, high-pressure liquid chromatography.

0006-2960/87/0426-3303$01.50/0