Pless, R. C., Levitt, L. M., & Bessman, M. J. (1981) *Biochemistry 20*, 6235.
Que, B. G., Downey, K. M., & So, A. G. (1978) *Biochemistry 17*, 1603.
Que, B. G., Downey, K. M., & So, A. G. (1979) *Biochemistry 18*, 2064.
Randerath, K., & Randerath, E. (1964) *J. Chromatogr. 16*, 111.
Wang, T. S.-F., Sedwick, W. D., & Korn, D. (1974) *J. Biol. Chem. 249*, 841.
Watanabe, S. M., & Goodman, M. F. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6429.

β Subunit of Rat Liver Mitochondrial ATP Synthase: cDNA Cloning, Amino Acid Sequence, Expression in *Escherichia coli*, and Structural Relationship to Adenylate Kinase†

David N. Garboczi, Arthur H. Fox, Sandra L. Gerring, and Peter L. Pedersen*
Laboratory for Molecular and Cellular Bioenergetics, Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received June 16, 1987; Revised Manuscript Received September 3, 1987

ABSTRACT: The amino acid sequence of all but a few N-terminal residues of the β subunit of rat liver ATP synthase has been determined from cDNA clones. Rat liver F_1-β is shown to contain 17 amino acid differences from that reported for F_1-β of bovine heart, 2 differences of which involve differences in charge. This may account in part for the observation that bovine heart F_1 binds nucleotides with much greater affinity than the rat liver enzyme. Rat liver F_1-β also contains homologous regions with another nucleotide binding protein, adenylate kinase, for which high-resolution structural studies are available. Adjacent to one of these homologous regions is an eight amino acid stretch which bears striking homology to the phosphorylation region of the (Na^+,K^+)-ATPase. The combination of these two homology regions may constitute at least part of a nucleotide binding domain in F_1-β. Significantly, both rat liver and bovine heart β contain these regions of homology, whereas the 17 amino acid differences between the two enzymes lie outside this region. The possibility of a second nucleotide binding domain which differs between the two enzymes is discussed. A cDNA clone containing all the regions of homology as well as 11 of the 17 amino acid differences between the bovine heart and rat liver β subunits has been ligated into the bacterial expression vector pKK223-3. After transformation of a protease-deficient strain of *Escherichia coli*, this cDNA clone is expressed as a 36-kilodalton protein. Finally, further cDNA library screening and primer extension analysis using several oligonucleotide probes generated cDNAs which always terminated prior to that region of the β gene coding for the N-terminal region. A strong secondary structure of regulatory significance may, therefore, be a unique characteristic of the rat liver mRNA coding for the F_1-β subunit in this region.

The ATP synthase of rat liver mitochondria is responsible for the production of ATP from the respiration-derived electrochemical gradient across the mitochondrial inner membrane. The enzyme complex is composed of two parts, F_1 and F_0, each being a multisubunit complex. The hydrophilic and membrane-extrinsic F_1 complex contains the catalytic regions of the molecule and has been shown to bind and hydrolyze nucleotides [for recent reviews, see Cross (1981), Senior and Wise (1983), Wang (1983), Amzel and Pedersen (1983), and Hatefi (1985)]. The enzyme subunit structure is highly conserved throughout evolution and, from bacteria to mammals, is generally agreed to be composed of five different subunit types in the ratio of $\alpha_3\beta_3\gamma\delta\epsilon$ (Senior & Brooks, 1971; Catterall & Pedersen, 1971; Catterall et al., 1973; Esch & Allison, 1979). The primary sequences of the three largest subunits, α, β, and γ, are also highly conserved (Walker et al., 1982, 1985), and, intriguingly, significant homology exists between the β subunit and numerous other nucleotide binding proteins (Walker et al., 1982; Fry et al., 1985). On the basis of these homologies and on chemical modification studies of β using nucleotide analogues, the suggestion has been raised that the mechanism of nucleotide binding may have similar themes in differing proteins.

Experiments described here were undertaken with a 3-fold purpose in mind. First, obtaining the predicted amino acid sequence of the β subunit will aid in understanding the characteristic nucleotide binding of the rat liver F_1-ATPase. Reports from this and other laboratories indicate that the F_1-ATPase from rat liver binds nucleotides with much less affinity than does the bovine heart enzyme (Catterall & Pedersen, 1972; Cross & Nalin, 1982). Moreover, a maximum of only four nucleotide binding sites are readily detected in the rat liver enzyme (Williams et al., 1987), whereas six sites are readily detected in similar binding assays in the bovine heart enzyme (Cross & Nalin, 1982). These differences between closely related enzymes may arise from primary sequence differences. Second, the acquisition of the F_1-β amino acid sequence will aid and complement the X-ray crystallographic structure determination currently in progress. A 9-Å map of rat liver F_1 has been obtained (Amzel et al., 1982) and is currently being extended with data up to a resolution of 3.5 Å. Third, obtaining a cDNA clone and the development of

†Supported by NIH Grant CA 10951 to P.L.P.

0006-2960/88/0427-0553$01.50/0

an expression system for the rat liver β subunit are prerequisites for a program of mutagenesis and studies of β proteins selectively mutated. Homologies between various β subunits and among other nucleotide proteins suggest possible experiments involved with determining which amino acids are essential for β-subunit function.

Here we describe the selection and sequencing of cDNA clones for the β subunit of the rat liver F_1-ATPase, using antisera raised to purified and intact F_1-ATPase. The DNA sequence predicts a β subunit closely homologous but with several notable differences from that of bovine heart. The expression in *Escherichia coli* of one of the β clones selected here is also described.

Experimental Procedures

Materials

Restriction and DNA-modifying enzymes were obtained from New England Biolabs, United States Biochemicals, Bethesda Research Laboratories (BRL), Pharmacia, and Boehringer-Mannheim Biochemicals and were used according to the suppliers' instructions. Agarose used was Seakem GTG from FMC Corp. Microtiter plates (Immulon 2) were from Dynatech Laboratories. Guanidine isothiocyanate was from Fluka Chemical Corp. Acrylamide gel reagents were from Bio-Rad. Laboratory chemicals were from Sigma and Baker. M13mp18 and -19, pUC18 and -19, universal sequencing 17-mer and deletion subcloning primers, and oligo(dT)–cellulose were from New England Biolabs. Bacterial strains Y1090 and JM101 were gifts of Dr. D. W. Cleveland (The Johns Hopkins School of Medicine). Preparation of the 41-mer oligonucleotide 5′-CACAACGGCGCCGATGACTGCCACAATTTGCCCGGTGGCGG-3′ was by The JHU Biopolymers Institute, The Johns Hopkins School of Medicine. Random hexamers and pKK223-3 were obtained from Pharmacia. [^{35}S]dATP was from New England Nuclear. ^{32}P-Labeled nucleotides were from ICN. NA45 nitrocellulose (DEAE) membrane was obtained from Schleicher & Schuell. Rat liver expression cDNA libraries were the gifts of Drs. P. Churchill, S. Churchill, and S. Fleischer (Vanderbilt University) and of Dr. R. Hynes (Massachusetts Institute of Technology). A human F_1-ATPase β-subunit cDNA was the gift of Drs. S. Ohta and Y. Kagawa (Jichi Medical School, Japan). *E. coli* strain CH1959 (CGSC 6790) was obtained from the *E. coli* genetic stock center at the Yale University School of Medicine. A sequence analysis program (DPSA) was obtained from Dr. Ch. Marck (Cente d'Etudes Nucleaires de Saclay, France) (Marck, 1986).

Methods

Antiserum Production and Characterization. F_1-ATPase purified by the method of Catterall and Pedersen (1971) was emulsified with an equal volume of adjuvant and injected at multiple subcutaneous sites in female New Zealand white rabbits. The adjuvant for the first injection was Freund's complete adjuvant; subsequent injections contained Freund's incomplete adjuvant. Anti-F_1 antibody production was monitored by a radioimmunoassay using antigen adsorbed to microtiter wells and ^{125}I-labeled protein A to assess antibody binding. Immunoblots were performed essentially by the method of Towbin et al. (1979). Antibody binding to blots was visualized by protein A and autoradiography.

Library Screening. cDNA library screening with an anti-F_1 antiserum was performed by the method of Young and Davis (1983) using ^{125}I-labeled protein A to locate antibody binding to plaques on a lawn of bacterial strain Y1090. Blocking of nonspecific sites on the filters, dilution of antiserum and of protein A, and washing of filters used a buffer containing 150 mM NaCl, 10 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl),[1] pH 7.4, and 0.1% Brij-58. Screening with a [^{32}P]dATP-labeled DNA fragment was performed by the method of Benton and Davis (1977). DNA fragments were labeled by the random primer method using the Klenow fragment of DNA polymerase I (Feinberg & Vogelstein, 1983). Positive plaques were purified by three further cycles of plating and screening at lower phage densities. Phage DNA was isolated essentially as described in Maniatis et al. (1982).

Northern Blot Analysis. Total RNA was isolated from rat liver by the guanidinium thiocyanate procedure of Chirgwin et al. (1979). RNA was enriched for poly(A)-containing sequences by two passages over oligo(dT)–cellulose (Aviv & Leder, 1972). RNA eluted from the column was fractionated by electrophoresis through 1.4% agarose gels containing formaldehyde (Goldberg, 1980) and transferred to nitrocellulose filters. Filters were hybridized with DNA labeled by the random primer method using [^{32}P]dATP or by polynucleotide kinase using [^{32}P]ATP. Filters were exposed to X-ray film at −70 °C with an intensifying screen.

Analysis of Clones. Restriction digests of recombinant phage DNA were analyzed by staining with ethidium bromide after electrophoresis on 0.8% agarose gels and by hybridization to fragments transferred to nitrocellulose filters. The filters were hybridized with a ^{32}P-labeled cDNA, washed to moderate stringency, and subjected to autoradiography at −70 °C with an intensifying screen. cDNA inserts for DNA sequencing were separated on agarose gels, recovered from the gels by using a DEAE membrane according to the manufacturer's instructions, and ligated into the *Eco*RI site of M13mp19. M13 phage were grown on bacterial strain JM101, and sequencing template DNA was purified according to the procedure in the BRL M13 Cloning/Dideoxy Sequencing Manual (Bethesda Research Laboratories, Gaithersburg, MD). Sequence determination was made according to the dideoxy chain termination procedure (Sanger et al., 1977) using a universal 17-mer sequencing primer and [^{35}S]dATP (Biggin et al., 1983). Sequential M13 subclones for sequencing were prepared by deletion of single-stranded phage DNA with T4 DNA polymerase by the method of Dale et al. (1985). Both strands and each nucleotide were sequenced approximately twice.

Expression of cDNA in E. coli. The C4 cDNA was ligated into the *Eco*RI site of the expression vector pKK223-3 which directs expression of inserted sequences under the control of the IPTG-inducible *trp-lac* (*tac*) promoter. Proper orientation of the insert was determined by restriction digests of plasmid DNA prepared by the boiling method (Maniatis et al., 1982). The expression plasmid was purified on a CsCl gradient and transformed into *E. coli* strain CH1959 which lacks the *lon* protease. Single colonies of transformants were picked and grown in minimal medium overnight. An aliquot of this overnight culture was used to inoculate 50 mL of LB medium which was shaken at 37 °C. At an OD_{600} = 0.7, IPTG was added to 1 mM, and shaking was continued for 1 h. At three time points, equal OD_{600} samples were sedimented, resuspended in Laemmli sample buffer, frozen/thawed with liquid nitrogen, heated at 100 °C for 2 min, forced through a 23-gauge needle, and loaded on 12% SDS–polyacrylamide gels

[1] Abbreviations: Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; IPTG, isopropyl thiogalactopyranoside; SDS, sodium dodecyl sulfate; bp, base pair(s); kDa, kilodalton(s); NMR, nuclear magnetic resonance; kb, kilobase(s); AMP-PNP, 5′-adenylyl imidodiphosphate; NBD-Cl, 7-chloro-4-nitro-2,1,3-benzoxadiazole.

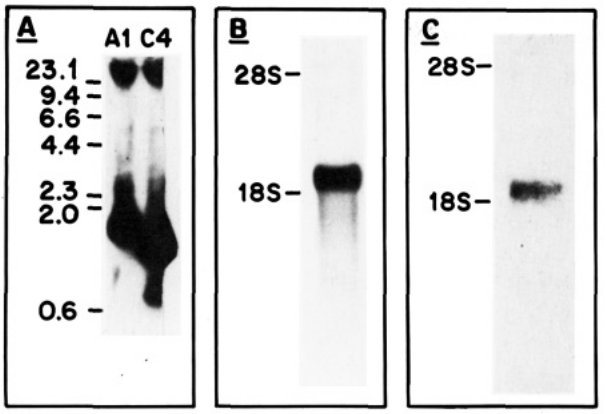

FIGURE 1: Identity of cDNA and size of mRNA of the rat liver F_1-β subunit determined by hybridizations. (A) Southern blot of *Eco*RI digest of cDNA clones A1 and C4 separated on 0.8% agarose showing hybridization of A1 insert at 1700 bp and of C4 insert at 1400 bp. Size markers are the fragments sized in kilobase pairs of a *Hin*dIII digest of λ DNA. Hybridization conditions: 50% formamide, 42 °C, 16 h. Washing conditions: 1× SSC (standard saline citrate, 20× SSC is 3 M sodium chloride and 0.3 M sodium citrate, pH 7.0), 65 °C, 0.1% SDS, 2 h. At this stringency, additional hybridization is visible at 870 bp in the C4 lane. Restriction mapping of C4 positions the 870 bp fragment 3′ to the poly(A) region of C4 (results not shown), thereby indicating that its hybridization is fortuitous and that it is unlikely to be related to the β subunit. (B) Autoradiogram of hybridization of labeled A1 cDNA to rat liver poly(A) containing RNA. Hybridization is to a band of 2100 bp. (C) Autoradiogram of end-labeled 41-mer oligonucleotide to poly(A) containing rat liver RNA. Hybridization is to a band of 2100 bp.

(Laemmli, 1970). Duplicate gels were loaded for Coomassie staining and for immunoblotting with anti-F_1 antiserum.

RESULTS

In preparation for selecting cDNA clones from a rat liver library, we raised an antiserum in rabbits to the rat liver F_1-ATPase complex. The antiserum reacted with the α, β, and γ subunits on protein blots of F_1 and only with those same subunit bands on blots of total proteins of rat liver cells. This antiserum was used to screen a rat liver cDNA library in the expression vector λ gt11. The first-round screening of 450 000 phage resulted in 11 positively signaling plaques. Of these 11 plaques, we purified as single plaques 4 phage that reacted with the antiserum on each of 3 successive rounds of screening. Purified and *Eco*RI-restricted DNAs from these clones were electrophoresed on agarose gels, revealing four differently sized inserts, two of which appeared to be long enough to encode about two-thirds of the β-subunit sequence.

Since antibody screening of expression libraries can often select clones containing cross-reacting epitopes but of no other relationship to the original antigen, we confirmed the identity of the clones by their hybridization with a human β-subunit cDNA. As is shown in Figure 1A, the human clone did hybridize under moderately stringent conditions to two putative rat liver β clones that were digested with *Eco*RI and blotted to filters. Hybridization signals are seen at 1700 bp in clone A1 and at 1400 bp in clone C4. This result showed that at least two of our clones indeed contained sequences of the rat liver β subunit. To determine the length and number of the β-subunit mRNA(s), we labeled clone A1 and hybridized it to a blot of poly(A) containing rat liver RNA. As is shown in Figure 1B, hybridization is seen at a band of about 2.1 kilobases in length. DNA sequencing of the A1 clone revealed the presence of insert sequences 3′ to the poly(A) region of the β-subunit cDNA. These sequences are presumed to be a cloning artifact introduced during the construction of the library, when two cDNAs were ligated into a single vector molecule without an intact *Eco*RI site between them. To ensure that these unidentified sequences were not hybridizing on the Northern blot, we probed a second poly(A) RNA blot with a labeled synthetic oligomer consisting of 41 bases located in the 5′-most β sequence we had obtained (see Figure 2 for the position of the oligomer in the sequence). In Figure 1C, a 2.1-kb band is identified by its hybridization to the oligomer. This result confirms that the 2.1-kb band is the β-subunit mRNA. This also indicates that there is likely to be only one mRNA species for the β subunit, although more than one species that are similar in size is possible.

Since neither of the clones determined to be rat liver β subunit by their hybridization to the human probe appeared to be long enough to account for the 2.1-kb mRNA seen on the Northern blot, we rescreened both libraries looking for more β sequence. Using a labeled restriction fragment of the human cDNA as a probe, we selected three additional clones from a second rat liver cDNA library in λ gt11. The probe was an *Eco*RI–*Sty*I fragment extending from the 5′ *Eco*RI site to the *Sty*I site located 662 bp away in the human cDNA and included the 5′ untranslated region of the human cDNA. The longest insert containing new sequence found with this probe was one of approximately 457 bp which overlapped approximately 100 bases of the longest antibody-selected clone (C4). Upon excision from the phage, this insert was cut into two fragments by *Eco*RI. The fragment that represented additional 5′ β sequence was 357 bp in length. Preliminary sequencing of this clone and of C4 showed that the β cDNA contains an internal *Eco*RI site which had not been methylated in the majority of the clones isolated from the first cDNA library. Complete DNA sequencing of the 357 bp insert and of the C4 insert was performed by using the deletion subcloning strategy of Dale et al. (1985). Subcloning of single-stranded M13 by this method enabled a rapid and ordered sequencing of the 1220 bp C4 cDNA. Each nucleotide on both strands was sequenced approximately twice.

In Figure 2, the rat liver cDNA sequence and the predicted amino acid sequence are shown. Comparison with the other mammalian β-subunit sequences reveals that the clones account for the entire coding region of the mRNA except for seven amino acids at the N-terminal. These were not encoded by any of the clones obtained. The clones include a 3′ untranslated region of 149 bp. The DNA sequence and the predicted amino acid sequence confirmed the clones' identity as being of the β-subunit sequence. At the nucleotide level, the coding regions of the rat liver cDNA and the human cDNA are 90% homologous, differing in 136 out of 1416 nucleotides. At the amino acid level the conservation of sequence is greater than 95%. The differences in amino acids between the rat liver β subunit and the β subunits of human and bovine are summarized in Table I. The predicted rat liver β subunit differs by 17 residues from the amino acid sequence of the bovine heart β subunit. Two of these differences involve charged residues giving the rat liver β subunit two net additional negative charges. Rat liver and human β-subunit amino acid sequences differ by seven residues.

Expression experiments were performed as described under Methods. The C4 clone was inserted into a vector which enables the expression of an insert under the inducible control of a hybrid *E. coli* promoter. The C4 clone codes for a protein of 353 amino acids with a molecular weight of approximately 36 000. We transformed the plasmid into a protease-deficient (*lon*$^-$) strain of *E. coli*. Figure 3 shows an immunoblot of total proteins from bacteria containing the original plasmid and

```
1/1                                   31/11                                 61/21
GAA TTC CGC ACC GCC ACC GGG CAA ATT GTG GCA GTC ATC GGC GCC GTT GTG GAT GTC CAG TTC GAT GAG GGA TTA CCA CCT ATC CTA AAT
            thr ala thr gly gln ile val ala val ile gly ala val val asp val gln phe asp glu gly leu pro pro ile leu asn

91/31                                 121/41                                151/51
GCC TTG GAA GTG CAA GGC AGG GAG AGC AGA CTG GTT TTG GAG GTA GCC CAG CAT TTA GGG GAG AGC ACC GTC AGA ACT ATT GCT ATG GAT
ala leu glu val gln gly arg glu ser arg leu val leu glu val ala gln his leu gly glu ser thr val arg thr ile ala met asp

181/61                                211/71                                241/81
GGC ACT GAA GGC TTG GTT AGA GGC CAG AAA GTA CTG GAT TCG GGG GCA CCA ATC AAA ATT CCT GTT GGT CCT GAG ACC TTG GGC AGA ATC
gly thr glu gly leu val arg gly gln lys val leu asp ser gly ala pro ile lys ile pro val gly pro glu thr leu gly arg ile

271/91                                301/101                               331/111
ATG AAT GTC ATT GGA GAA CCT ATT GAT GAG AGA GGT CCT ATC AAA ACC AAA CAA TTC GCT CCT ATT CAT GCT GAG GCT CCT GAA TTC ATA
met asn val ile gly glu pro ile asp glu arg gly pro ile lys thr lys gln phe ala pro ile his ala glu ala pro glu phe ile

361/121                               391/131                               421/141
GAG ATG AGT GTT GAA CAG GAA ATT CTG GTG ACT GGT ATA AAG GTT GTG GAT CTG CTG GCC CCA TAC GCC AAG GGT GGG AAA ATC GGA CTC
glu met ser val glu gln glu ile leu val thr gly ile lys val val asp leu leu ala pro tyr ala lys gly gly lys ile gly leu

451/151                               481/161                               511/171
TTC GGA GGT GCT GGT GTT GGA AAG ACA GTA CTG ATC ATG GAG CTA ATC AAC AAT GTT GCT AAA GCC CAT GGT GGT TAT TCT GTA TTT GCT
phe gly gly ala gly val gly lys thr val leu ile met glu leu ile asn asn val ala lys ala his gly gly tyr ser val phe ala

541/181                               571/191                               601/201
GGT GTT GGT GAG AGG ACC CGT GAG GGC AAT GAT TTA TAC CAT GAA ATG ATT GAG TCT GGT GTT ATC AAC CTA AAA GAT GCC ACT TCC AAG
gly val gly glu arg thr arg glu gly asn asp leu tyr his glu met ile glu ser gly val ile asn leu lys asp ala thr ser lys

631/211                               661/221                               691/231
GTA GCG TTG GTA TAT GGG CAG ATG AAT GAA CCG CCT GGT GCT CGT GCC CGG GTA GCT CTG ACT GGT CTG ACT GTT GCT GAA TAC TTC AGA
val ala leu val tyr gly gln met asn glu pro pro gly ala arg ala arg val ala leu thr gly leu thr val ala glu tyr phe arg

721/241                               751/251                               781/261
GAC CAG GAA GGC CAA GAT GTC CTG CTG TTT ATT GAC AAC ATC TTC CGC TTC ACC CAG GCT GGC TCA GAG GTA TCT GCC TTA TTG GGC AGG
asp gln glu gly gln asp val leu leu phe ile asp asn ile phe arg phe thr gln ala gly ser glu val ser ala leu leu gly arg

811/271                               841/281                               871/291
ATC CCG TCT GCT GTA GGC TAC CAG CCT ACC CTA GCC ACT GAC ATG GGT ACA ATG CAG GAA AGA ATC ACC ACC ACC AAG AAG GGC TCG ATC
ile pro ser ala val gly tyr gln pro thr leu ala thr asp met gly thr met gln glu arg ile thr thr thr lys lys gly ser ile

901/301                               931/311                               961/321
ACC TCA GTG CAG GCT ATC TAT GTG CCA GCT GAT GAC CTG ACT GAC CCT GCC CCT GCA ACT ACC TTT GCC CAT TTG GAT GCT ACT ACT GTG
thr ser val gln ala ile tyr val pro ala asp asp leu thr asp pro ala pro ala thr thr phe ala his leu asp ala thr thr val

991/331                               1021/341                              1051/351
CTG TCC CGT GCT ATT GCT GAG TTG GGC ATC TAT CCA GCT GTG GAT CCG CTG GAC TCC ACC TCT CGA ATT ATG GAT CCC AAC ATC GTT GGC
leu ser arg ala ile ala glu leu gly ile tyr pro ala val asp pro leu asp ser thr ser arg ile met asp pro asn ile val gly

1081/361                              1111/371                              1141/381
AGT GAG CAT TAT GAT GTT GCT CGT GGG GTG CAA AAG ATC CTG CAG GAC TAC AAA TCT CTC CAG GAC ATC ATT GCC ATC TTG GGT ATG GAT
ser glu his tyr asp val ala arg gly val gln lys ile leu gln asp tyr lys ser leu gln asp ile ile ala ile leu gly met asp

1171/391                              1201/401                              1231/411
GAA CTT TCT GAG GAA GAT AAA TTG ACT GTG TCC AGG GCA AGG AAG ATA CAG CGC TTC TTG TCA CAG CCA TTC CAG GTT GCT GAG GTC TTC
glu leu ser glu glu asp lys leu thr val ser arg ala arg lys ile gln arg phe leu ser gln pro phe gln val ala glu val phe

1261/421                              1291/431                              1321/441
ACA GGT CAC ATG GGA AAG CTG GTG CCC CTG AAG GAG ACC ATT AAA GGA TTC CAG CAG ATC TTA GCA GGT GAC TAT GAC CAT CTC CCG GAA
thr gly his met gly lys leu val pro leu lys glu thr ile lys gly phe gln gln ile leu ala gly asp tyr asp his leu pro glu

1351/451                              1381/461                              1411/471
CAA GCC TTC TAC ATG GTG GGA CCC ATT GAA GAA GCT GTG GCA AAG GCT GAC AAG CTG GCA GAG GAG CAT GGG TCG TGA GGG GCC CTT CAG
gln ala phe tyr met val gly pro ile glu glu ala val ala lys ala asp lys leu ala glu glu his gly ser

1441/481                              1471/491                              1501/501
CCA AAC ACA ACA GCA CTC TGC ACT GAC CTC CAT GCT GAG AGC TCA GTT TGC CAT GTA GGC CAC ACA AGA GCC TTG ATT GAA GAT GTG ATG

1531/511                              1561/521
TTC TCT CTG AAG AGT ATT TAA AGT TTT CAA TAA AGT ATA TAC CCT
```

FIGURE 2: cDNA and deduced amino acid sequence of the β subunit. The 1577 bp and the 472 predicted amino acids of the cDNAs are listed. The rat liver β-subunit sequence begins at the fourth codon as the first three codons are a *Eco*RI linker introduced during library construction. Underlining of nucleotides denotes the position of the 41-mer oligonucleotide (see text). Underlining of amino acids denotes homologous residues shared between the β subunit and adenylate kinase. Sequence numbering is four positions different from the bovine β subunit (rat liver residue number + 4 = bovine residue number). Ser-475 is the last amino acid followed by a termination codon and a 149 bp 3′ untranslated region. Not shown here is the 45 bp poly(dA) region found on the C4 clone. (Note: As emphasized in the text, a small portion of the N-terminal region, predicted to be seven amino acids, was not encoded by the clones obtained.)

bacteria containing the recombinant C4-containing plasmid. The blot was incubated with the anti-F_1 serum followed by ^{125}I-labeled protein A. One hour after induction of the cells, a labeled band of approximately 36 000 daltons is seen, indicating that the rat liver β cDNA clone C4 is being expressed. This band was not seen in protein blots of *lon*$^+$ *E. coli* strains containing the original plasmid.

Discussion

Results reported in this paper detail the molecular cloning of overlapping cDNAs for the β subunit of the rat liver F_1-ATPase. DNA sequencing of these clones by a newly described deletion subcloning method (Dale et al., 1985) and the Sanger chain termination method showed that the clones represent 472 amino acids of the β subunit. An unknown number of amino acids, predicted to be seven by comparison with other β subunits, was not encoded by the clones obtained.

Significantly, the rat liver β-subunit sequence differs in 17 residues from that of bovine heart and in 7 residues from the predicted sequence of human (HeLa cell) β subunit (Table I). There are a number of β-subunit sequences known (>12) of organisms from bacteria to mammals, and in all species studied, there is high conservation of sequence (Walker et al., 1985). Although this conservation most likely reflects a generally conserved mechanism of ATP synthesis and ATP hydrolysis by the F_0F_1 complex in many species, the β-subunit differences should not be ignored. It is likely that they play some role either in governing the catalytic quality of ATP

Table I: Amino Acid Differences among Mammalian β Subunits

residue[a]	rat liver	bovine[b]	human[c]
8	Thr	Ala	Ala
9	Ala	Thr	Ala
12	Gln	Arg	Arg
43	Ser	Thr	Thr
83	Lys	Arg	Lys
115	Pro	Ala	Pro
124	Ile	Val	Met
132	Ile	Leu	Ile
137	Ile	Asp	Ile
165	Leu	Phe	Leu
224	Glu	Gln	Gln
288	Asp	Asn	Asp
324	Thr	Val	Thr
359	Asp	Asn	Asp
428	Met	Leu	Met
448	Asp	Glu	Glu
478	Gly	Ser	Ser

[a] Sequence numbering from Walker et al. (1983). [b] Sequence from Walker et al. (1983). [c] Sequence from Ohta and Kagawa (1986).

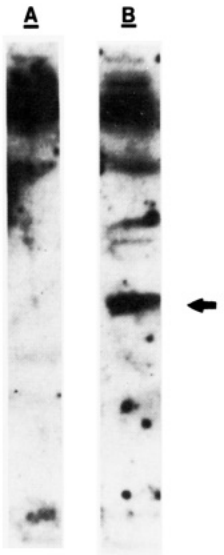

FIGURE 3: Immunoblot of cDNA protein expressed in *E. coli*. Total *E. coli* proteins were obtained and immunoblots performed after 1 h of induction as described under Methods. (Lane A) Blot of proteins from bacteria containing the expression plasmid without the C4 insert. (Lane B) Blot of proteins from bacteria containing the expression plasmid with the C4 insert.

synthase complexes (i.e., their turnover number) or in regulating their activity. Along these lines, it should be noted that two of the differences elucidated here between rat liver and bovine heart β's are differences in charged residues (Table I). These two, and other sequence differences, may explain the sharp contrast in the binding of nucleotides between the rat and bovine enzymes both in the number and in the affinity of binding sites (Catterall & Pedersen, 1972; Cross & Nalin, 1982; Williams et al., 1987). Only four nucleotides readily bind to 1 mol of rat liver F_1 (Williams et al., 1987) whereas six are readily detected in binding assays on the bovine heart enzyme (Cross & Nalin, 1982). In addition, the heart enzyme is reported to bind AMP-PNP with K_d values as low as 0.1 μM (Cross & Nalin, 1982) whereas the lowest values obtained for the liver enzyme are about 1 μM (Williams et al., 1987).

Two other features of the predicted amino acid sequence of the rat liver β subunit are noteworthy and merit discussion. The first is the β-subunit sequence homology to the active-site

Table II: Three Regions of Sequence Homology between Rat Liver β Subunit and Adenylate Kinase[a]

```
I)
F1-β   (156-177)  G-G-A-G-V-G-K-T-V-L-I-M-E-L-I-N-N-V-A-K-A-H
                  | |   |   | |               | |           |
AdK    (15-36)   G-G-P-G-S-G-K-G-T-Q-C-E-K-I-V-H-K-Y-G-Y-T-H

II)
F1-β   (248-256)  G-Q-D-V-L-L---F-I-D
                  |     | |     | |
AdK    (110-119)  A-Q-P-T-L-L-L-Y-V-D

III)
F1-β   (311-333) Y-V-P-A-D-D-L-T-D-P-A-P-A-T-T-F-A-H-L-D-A-T-T
                   |   | |       |               | | | |
AdK    (172-194) K-V-N-A-E-G-S-V-D-N-V-F-S-Q-V-C-T-H-L-D-A-L-K
```

[a] Adenylate kinase sequence from Kuby et al. (1984).

Table III: Sequence Homology between F_0F_1 and E_1E_2 ATPases

Rat liver β subunit	(298-305)	T-T -K-K[d]-G-S-I[d]-T
E. coli β subunit[a]	(285-292)	S-T -K-T -G-S-I- T
Na^+,K^+-ATPase[b]	(368-375)	S-D[e]-K- T-G-T-L -T
Ca^{++}-ATPase[c]	(350-357)	S-D[e]-K- T-G-T-L- T

[a] Sequence from Walker et al. (1984). [b] Sequence from Shull et al. (1985). [c] Sequence from MacLennan et al. (1985). [d] Lysine that labels with 8-azido-ATP (Hollemans et al., 1983). [e] Aspartic acid which is phosphorylated during catalysis (MacLennan et al., 1985; Shull et al., 1985).

region of adenylate kinase (Walker et al., 1982; Fry et al., 1985), and the second is the homology to the catalytic region of ATPases like the (Na^+/K^+)- and Ca^{2+}-ATPases (Ernster et al., 1986). In Table II, three areas of homology between the rat liver β subunit and adenylate kinase are shown. X-ray structural studies (Pai et al., 1977) and high-resolution NMR studies (Fry et al., 1985, 1987) of adenylate kinase have provided a wealth of available structural data which have allowed predictions to be made about the functions of homologous residues in the β subunit (Walker et al., 1982; Fry et al., 1986; Duncan et al., 1986). On the basis of the homologies noted here, these predictions may also be applied to the rat liver β subunit.

The predicted secondary structure of a portion of the rat liver β subunit which is the size of adenylate kinase and which was chosen by aligning the two proteins by their homologous regions is shown in Figure 4. Alternating β sheets and α helices interspersed with turns are predicted, and, interestingly, these secondary structural elements can be folded into a β–α–β structure. The central feature of such a structure is the five-stranded parallel β sheet that is a common property of nucleotide binding folds (Richardson, 1981). A drawing of adenylate kinase based on its X-ray crystal structure showing such a parallel β sheet is presented in Figure 5A.

On the basis of predicted secondary structures and on homology with adenylate kinase, Figure 5B shows a model for the folding of the rat liver β-subunit protein. The areas of homology between the β subunit and adenylate kinase are indicated in Figure 5A,B. As is also indicated in Figure 5B, some β-subunit amino acid residues, surmised from chemical modification studies to be of catalytic importance, assume locations proximal to the predicted nucleotide binding site. Recently, a similar model for the folding of the *E. coli* β subunit was presented (Duncan et al., 1986). As the mammalian and bacterial β subunits are 80% homologous at the amino acid level in the region of homology, it is probable that both models are describing very similar structures.

A short segment of the rat liver β-subunit amino acid sequence shows homology to the phosphorylation region of the (Na^+/K^+)- and Ca^{2+}-ATPases (Pedersen & Carafoli, 1987)

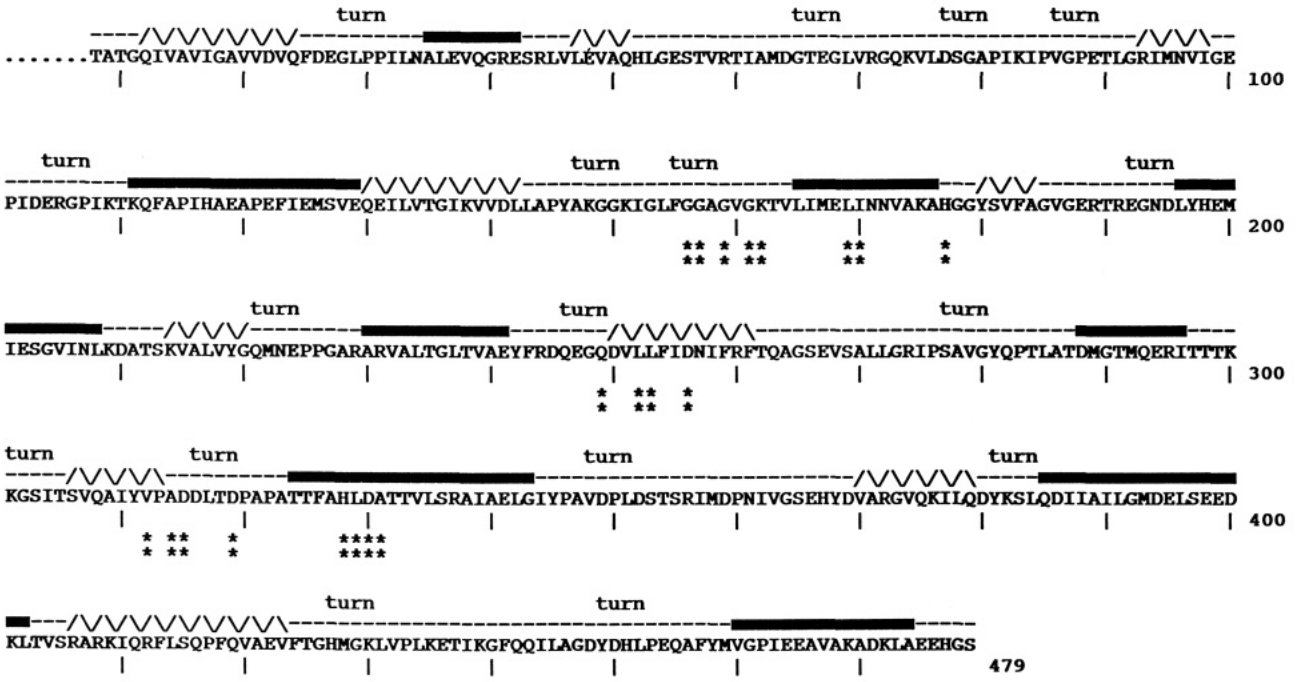

FIGURE 4: Predicted secondary structure of the rat liver β subunit. Above the single-letter form of the predicted rat liver β-subunit amino acid sequence are secondary structural features as predicted by the method of Chou and Fasman (1978). (■) α helix; (/\/\) β sheet; (turn) β turn; (stacked asterisks) homologous amino acids with adenylate kinase. Numbering according to Runswick and Walker (1983).

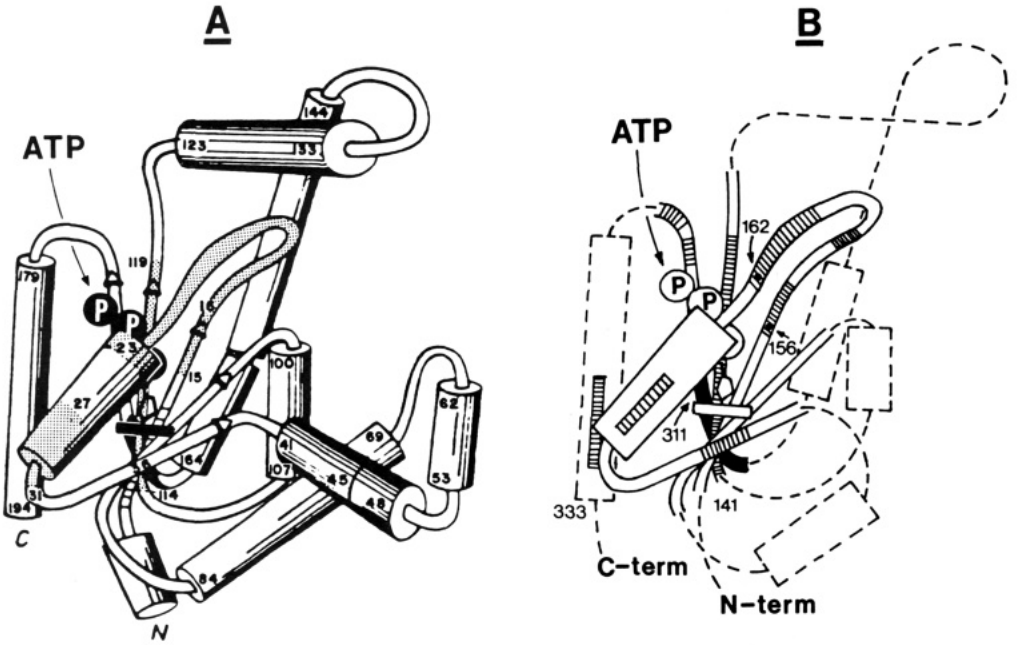

FIGURE 5: Adenylate kinase and a model for the nucleotide binding domain of the rat liver β subunit. (A) Drawing of the X-ray crystal structure of adenylate kinase (Pai et al., 1977; Fry et al., 1985). Note ATP in its binding site; the adenine is seen edge-on as a black rectangle, and the phosphates are black closed circles labeled "P". Homologous residues with several ATP binding proteins are shown by stippling (Fry et al., 1985). Sequence numbering is that of Kuby et al. (1984). (B) A speculative structure of the rat liver β subunit. The predicted secondary structures of amino acids numbered 141–333 of the β subunit can be folded into a five-stranded parallel β sheet connected by α helices analogous to adenylate kinase. Here the adenine is seen edge-on and is an open rectangle; the phosphates are depicted as open circles labeled "P". (||) Homologous amino acids with adenylate kinase; (■) homologous segment with (Na^+/K^+)- and Ca^{2+}-ATPases. Some residues of the β subunit that have been chemically modified are shown: Lys-162, labeled by NBD-Cl (Andrews et al., 1984b); Tyr-311, labeled by NBD-Cl (Andrews et al., 1984a). Gly-156 is predicted to be important for the functioning of the "flexible loop" (Fry et al., 1986).

(Table III) as reported recently (Ernster et al., 1986). This homology may indicate an evolutionary relationship of functional significance between the latter ATPases and the F_0F_1-ATPase. The marked aspartic acid residue in Table III is thought to be the aspartic acid that is known to be phosphorylated during the catalytic cycle of the (Na^+,K^+)-ATPase (Shull et al., 1985) and the Ca^{2+}-ATPase (MacLennan et al., 1985). This residue is not conserved in the F_1-ATPases but is a threonine residue. In bovine F_1-ATPase, it has been shown that 8-azido-ATP modifies a lysine (Hollemans et al., 1983), just two residues away from the threonine, suggesting the proximity of this region to a nucleotide binding region.

Recent reports from two different laboratories indicate that each β subunit may contain two nucleotide binding domains (Khanashvili & Gromet-Elhanan, 1985; Kironde & Cross, 1987). It seems possible that one of these, perhaps the one

predicted from homology arguments with adenylate kinase, represents the catalytic site, at least in part. The remaining part of the catalytic site may involve one or several nucleophilic residues like tyrosine-345 (Cross et al., 1987), which in the 3-dimensional structure of the β subunit may lie near the nucleotide binding domain predicted by comparison with adenylate kinase and other ATP binding proteins. A second site, perhaps regulatory or structural in function, may be less conserved and explain differences in nucleotide binding properties observed for F_1-ATPases from different sources. In this regard, it is important to note that the 17 amino acid differences observed here between the bovine heart and rat liver F_1 β subunits lie outside the nucleotide binding region predicted by comparison with adenylate kinase.

We expressed the partial β cDNA C4 in *E. coli* (Figure 3). This cDNA contains all of the adenylate kinase homology regions, and, therefore, the truncated β subunit it encodes may remain functional. This may enable nucleotide binding to be directly localized to a particular region of the β-subunit polypeptide. As the N-terminal region varies across species significantly more than the C-terminal region, the expression of this partial clone will allow the study of the highly conserved C-terminal region. Finally, the expression of the rat liver β subunit in bacteria is a first step in the direction of performing mutational studies on the protein. The specificity of modification through mutation will be a useful tool in the elucidation of the workings of this highly conserved enzyme.

Finally, we should note that we carried out a number of experiments in an attempt to elucidate both the N-terminal and regulatory region of the rat liver β subunit. These experiments included extensive screening of our cDNA libraries and primer extension analysis using several oligonucleotide probes. In all cases, sequences obtained terminated prior to the region of the β gene coding for the N-terminal region. Although these efforts may have been without success for technical reasons, we believe it more likely that a strong secondary structure of regulatory significance may be a unique characteristic of the rat liver mRNA coding for the F_1-β subunit in this region.

ACKNOWLEDGMENTS

We thank Drs. D. W. Cleveland, K. Sullivan, and T. Yen for advice on DNA methods. We acknowledge the generous gifts of cDNA libraries from Drs. P. Churchill, S. Churchill, S. Fleischer, and R. Hynes, a human β-subunit cDNA from Drs. S. Ohta and Y. Kagawa, and a sequence manipulation program for the Apple IIe from Dr. Ch. Marck. We are grateful to Drs. D. Fry, A. Mildvan, and D. Shortle for helpful discussions during the course of this work. We thank Drs. R. Kaplan, R. Nakashima, R. Pratt, and K. Arora and J. Barnard for critically reading the manuscript.

Registry No. ATP synthase, 37205-63-3; adenylate kinase, 9013-02-9.

REFERENCES

Amzel, L. M., & Pedersen, P. L. (1983) *Annu. Rev. Biochem. 52*, 801–824.

Amzel, L. M., McKinney, M., Narayanan, P., & Pedersen, P. L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5852–5856.

Andrews, W. W., Hill, F. C., & Allison, W. S. (1984a) *J. Biol. Chem. 259*, 8219–8225.

Andrews, W. W., Hill, F. C., & Allison, W. S. (1984b) *J. Biol. Chem. 259*, 14378–14382.

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Benton, W. P., & Davis, R. W. (1977) *Science (Washington, D.C.) 196*, 180–184.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Catterall, W. A., & Pedersen, P. L. (1971) *J. Biol. Chem. 246*, 4987–4994.

Catterall, W. A., & Pedersen, P. L. (1972) *J. Biol. Chem. 247*, 7969–7976.

Catterall, W. A., Coty, W. A., & Pedersen, P. L. (1973) *J. Biol. Chem. 248*, 7427–7431.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Cross, R. L. (1981) *Annu. Rev. Biochem. 50*, 681–714.

Cross, R. L., & Nalin, C. M. (1982) *J. Biol. Chem. 257*, 2874–2881.

Cross, R. L., Cunningham, D., Miller, C. G., Xue, Z., Zhou, J.-M., & Boyer, P. D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5715–5719.

Dale, R. M. K., McClure, B. A., & Houchins, J. P. (1985) *Plasmid 13*, 31–40.

Duncan, T. M., Parsonage, D., & Senior, A. E. (1986) *FEBS Lett. 208*, 1–6.

Ernster, T., Hundal, B., Sandri, G., Wojtczak, L., Grinkevich, V. A., Modyanov, N. N., & Ovchinnikov, Yu. A. (1986) *Chem. Scr. 26B*, 273–279.

Esch, F. S., & Allison, W. S. (1979) *J. Biol. Chem. 254*, 10740–10746.

Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.

Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1985) *Biochemistry 24*, 4680–4694.

Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 907–911.

Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1987) *Biochemistry 26*, 1645–1655.

Goldberg, D. A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5794–5798.

Hatefi, Y. (1985) *Annu. Rev. Biochem. 54*, 1015–1069.

Hollemans, M., Runswick, M. J., Fearnley, I. M., & Walker, J. E. (1983) *J. Biol. Chem. 258*, 9307–9313.

Khanashvili, D., & Gromet-Elhanan, Z. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1886–1890.

Kironde, F. A. S., & Cross, R. L. (1987) *J. Biol. Chem. 262*, 3488–3495.

Kuby, S. A., Palmieri, R. H., Frischat, A., Fischer, A. H., Wu, L. H., Maland, L., & Manship, M. (1984) *Biochemistry 23*, 2393–2399.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

MacLennan, D. H., Brandl, C. J., Korczak, B., & Green, N. M. (1985) *Nature (London) 316*, 696–700.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marck, Ch. (1986) *Nucleic Acids Res. 14*, 583–590.

Ohta, S., & Kagawa, Y. (1986) *J. Biochem. (Tokyo) 99*, 135–141.

Pai, E. F., Sachsenheimer, W., Schirmer, R. H., & Schultz, G. E. (1977) *J. Mol. Biol. 114*, 37–45.

Pedersen, P. L., & Carafoli, E. (1987) *Trends Biochem. Sci. (Pers. Ed.) 12*, 186–189.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Runswick, M. J., & Walker, J. E. (1983) *J. Biol. Chem. 258*, 3081–3089.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Senior, A. E., & Brooks, J. C. (1971) *FEBS Lett. 17*, 327-329.
Senior, A. E., & Wise, J. G. (1983) *J. Membr. Biol. 73*, 105-124.
Shull, G. E., Schwartz, A., & Lingrel, J. B. (1985) *Nature (London) 316*, 691-695.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350-4354.
Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982) *EMBO J. 1*, 945-951.
Walker, J. E., Saraste, M., & Gay, N. J. (1984) *Biochim. Biophys. Acta 768*, 164-200.
Walker, J. E., Fearnley, I. M., Gay, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M., & Tybulewicz, V. L. J. (1985) *J. Mol. Biol. 184*, 677-701.
Wang, J. H. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 21-24.
Williams, N., Hullihen, J., & Pedersen, P. L. (1987) *Biochemistry 26*, 162-169.
Young, R. A., & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194-1198.

Expression of the *Drosophila* Type II Topoisomerase Is Developmentally Regulated†

Robert Fairman and Douglas L. Brutlag*
Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305
Received August 6, 1987; Revised Manuscript Received October 6, 1987

ABSTRACT: The expression of the type II topoisomerase from *Drosophila melanogaster* was studied during development and in tissue culture cells. RNA blot and protein blot analyses using probes specific for *Drosophila* topoisomerase II show that the enzyme is developmentally regulated. Levels of both RNA transcript and protein appear highest during early embryogenesis and pupation, periods which are known to have the highest mitotic activity. Tissue culture analysis using *Drosophila* K_c cells supports these results as levels of topoisomerase II message are higher in rapidly dividing cells than in quiescent cells. Analysis of topoisomerase II levels in early embryos suggests that levels are adequately high for the enzyme to act in DNA replication or segregation at termination of replication. Apparent in vivo proteolysis of topoisomerase II is seen throughout the life cycle, in spite of careful precautions. Whether these proteolytic fragments are important in vivo is still uncertain.

Topoisomerases are enzymes that are found in both prokaryotes and eukaryotes and are involved in maintaining the topology of DNA [for recent reviews, see Wang (1985) and Vosberg (1985)]. There are two types of topoisomerases, defined by their mechanism of action. Type I topoisomerases make a transient nick in DNA, allowing the broken strand to swivel around the intact strand to change the twist of the DNA. Type II topoisomerases cause a double-strand break in DNA and allow a nearby helix to pass through the broken helix to change the writhe of the DNA. Most type II enzymes require ATP for their action. Both enzymes make a change in the net supercoiling of the DNA. In prokaryotes, topoisomerases regulate the supercoil density of the DNA. In *Escherichia coli*, the type I topoisomerase removes negative supercoils, and the type II topoisomerase (DNA gyrase) induces negative supercoils in DNA. In eukaryotes, no purified supercoiling activity has been identified; however, both type I and type II topoisomerases can remove supercoils. In fact, eukaryotic topoisomerases can remove both negative and positive supercoils.

Whereas the functions of the prokaryotic enzymes are fairly well understood, relatively little is known about the function(s) of the eukaryotic topoisomerases. Double mutants in yeast show that there is an absolute requirement for topoisomerase II. The topoisomerase II requirement appears to be involved in replication termination (segregation) and medial nuclear division (Goto & Wang, 1985; Holm et al., 1985; DiNardo et al., 1983; Uemura & Yanagida, 1984). A class of drugs known as epipodophylotoxins has been used to study the in vivo role of topoisomerase II by trapping topoisomerase II–DNA complexes via a reversible covalent intermediate. Nelson et al. (1986) used teniposide (an epipodophylotoxin) to show that, in mammalian cells, topoisomerase II is associated with nascent DNA fragments near DNA replication forks, and Heck and Earnshaw (1986), using immunofluorescence labeling, showed that the appearance of topoisomerase II exactly parallels the onset of DNA replication in chicken cell lines. In addition to an enzymatic role, topoisomerase II has also been suggested to play a structural role. Berrios et al. (1985) have suggested that topoisomerase II is a component of the nuclear matrix in *Drosophila*. Earnshaw and Heck (1985), Earnshaw et al. (1985), and Gasser et al. (1986) have shown that topoisomerase II acts as a chromosomal scaffold protein both in chicken and in HeLa cells.

Other work has suggested a role for DNA topoisomerases in regulating transcription. Ryoji and Worcel (1984, 1985) described two types of chromatin in *Xenopus* oocytes. These types were called "dynamic" and "static". Dynamic chromatin is found to be torsionally constrained and was shown to be the transcriptionally active form of chromatin. Later, Kmiec and Worcel (1985) found that TFIIIA, the positive transcription factor of the 5S RNA gene, is required to make and maintain dynamic chromatin and confers a gyrase-like activity on the

† This work was supported by a grant from the NIH (GM28079).
* Address correspondence to this author.

0006-2960/88/0427-0560$01.50/0