Repeated Consensus Sequence and Pseudopromoters in the Four Coordinately Regulated Tubulin Genes of *Chlamydomonas reinhardtii*

KAREN J. BRUNKE,^{1*} JAMES G. ANTHONY,^{1†} EDMUND J. STERNBERG,^{1‡} AND DONALD P. WEEKS²

The Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111,¹ and Zeecon Corporation, Palo Alto, California 94304²

Received 3 January 1984/Accepted 6 March 1984

The 5' coding and promoter regions of the four coordinately regulated tubulin genes of *Chlamydomonas reinhardtii* have been mapped and sequenced. DNA sequencing data shows that the predicted N-terminal amino acid sequences of *Chlamydomonas* α - and β -tubulins closely match that of tubulins of other eucaryotes. Within the α_1 - and α_2 -tubulin gene set and the β_1 - and β_2 -tubulin gene set, both nucleotide sequence and intron placement are highly conserved. Transcription initiation sites have been located by primer extension analysis at 140, 141, 159, and 132 base pairs upstream of the translation initiator codon for the α_1 -, α_2 -, β_1 -, and β_2 -tubulin genes, respectively. Among the structures with potential regulatory significance, the most striking is a 16-base-pair consensus sequence [GCTC(G/C)AAGGC(G/T)(G/C)C(A)(C/A)G] which is found in multiple copies immediately upstream of the TATA box in each of the four genes. An unexpected discovery is the presence of pseudopromoter regions in two of the transcribed tubulin genes. One pseudopromoter region is located 400 base pairs upstream of the authentic α_2 -tubulin gene promoter, whereas the other is located within the transcribed 5' noncoding region of the β_1 -tubulin gene.

Genes which are induced coordinately in response to a stimulus or a developmental cue and which are not tightly clustered must each possess a common feature which directs their synchronous response to the regulatory signal. In a number of eucaryotic systems, that common feature appears to be a short sequence of nucleotides in the 5' boundary region of the coordinately regulated genes (5, 10, 14, 17-19, 24, 25, 28, 45, 49, 50, 56). Comparison of several developmentally and hormonally controlled genes, such as the chorion genes of silkworms (28), the glucocorticoid-regulated genes of humans (14), mice (5), and rats (19), and the steroid-induced egg white protein genes of chickens (24), reveals that in addition to the usual RNA polymerase II recognition signals there are distinct sequence homologies in the 5' promoter region within each of these gene sets. In general, the putative regulatory elements within a gene set are similar but not identical, quite short (9 to 24 base pairs [bp]), and often present more than once in the upstream region. Statistical analysis, such as that described by Davidson et al. (17), demonstrates a strikingly low probability that the putative control elements are present by chance in the upstream region of each member of the coordinately regulated gene set.

More definitive evidence that small regions of nucleotide homology can suffice to allow coordinate regulation comes from studies of gene systems in which regulation of one or more genes within a set has been altered by mutation (45) and, more convincingly, from analysis of systems in which the promoter regions of isolated genes have been altered and then reintroduced by transformation into homologous cells (18, 25, 56) or heterologous (10, 50) cell cultures. Thus, in yeast cells (18, 25, 56) it has been shown that coinduction of a number of genes involved in general amino acid metabolism is mediated primarily through 9-bp (consensus) sequences located upstream of the transcription initiation site

(cap site) and Goldberg-Hogness (TATA) box in each gene. Likewise, transformation of monkey COS cells or *Xenopus* oocytes with genes under the control of promoters from the *Drosophila* heat shock protein genes has identified a 15-bp consensus sequence which apparently governs response to heat shock regulatory signals (10, 50). Despite such progress with certain gene sets in revealing the nucleotide elements that are responsible for coordinate transcription, it is clear that there are still only a limited number of examples upon which we may draw in formulating more general rules of how eucaryotic organisms bring about the simultaneous induction of unlinked genes. The studies reported here of the tubulin induction system in *Chlamydomonas reinhardtii*, an organism that has recently been reported to be amenable to transformation by exogenous DNA (53), indicate that this may be an additional system which can be exploited for studies of coordinate gene regulation.

The removal of flagella from *C. reinhardtii* triggers a rapid and massive induction of synthesis of tubulin (44, 54, 60) and other flagellar proteins (34, 52). Paralleling the abrupt rise and fall in the production of α - and β -tubulin subunits during the 2- to 3-h period of flagellar regeneration is a synchronous increase and decrease in the synthesis of four distinct tubulin mRNAs (12). The two α - and two β -tubulin mRNAs are produced in concert (12, 54) from a set of four genes which are not closely linked (12). To investigate possible mechanisms involved in the coordinate expression of these genes we have sequenced the upstream promoter and coding regions of all four genes. Our analysis has revealed that in addition to some of the usual regulatory signals found in other eucaryotic promoter regions, such as the TATA box and mRNA initiation site, a number of unique sequences and interesting structural features are present in the *C. reinhardtii* tubulin gene promoter regions. Perhaps the most intriguing observation is the finding of multiple copies of a 16-bp consensus sequence a short distance upstream from the TATA box in each α - and β -tubulin gene. In addition, we have discovered sequences in the 5' region of one α -tubulin gene and one β -tubulin gene which we have designated as pseudopromoters since they exhibit marked similarities to

* Corresponding author.

† Present address: Zeecon Corporation, Palo Alto, CA 94304.

‡ Present address: Smith Kline & French Laboratories, Swedeland, PA 19479.

the authentic promoter regions of the tubulin genes but are not used for mRNA synthesis during tubulin induction.

MATERIALS AND METHODS

Construction and isolation of chimeric plasmids containing 5' regions of the four tubulin genes of *C. reinhardi*. Genomic clones of the *C. reinhardi* tubulin genes prepared in the λ vector, charon 30, have been previously described (12). To obtain smaller subclones containing the 5' regions for each of the tubulin genes, chimeric plasmids were constructed by first cutting the DNA from genomic clones with restriction enzyme *Pst*I. For each gene, a *Pst*I fragment containing much of the 5' region, the entire noncoding region, and a portion of the coding region was then isolated and inserted into the *Pst*I site of the ampicillin resistance gene of the plasmid vector, pBR322. Plasmids were used to transform HB101-competent bacteria with selection for Amp^r Tet^r colonies. Alternatively, selected DNA fragments from the genomic tubulin clones were ligated into the polylinker region of the pBR322 derivative plasmid, pUC9, and transformed into JM83-competent bacteria with selection of colorless colonies in the presence of 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (Xgal) (43). Minipreps of plasmid DNA from the constructed subclones were analyzed for size and for restriction enzyme cutting sites to confirm both the actual DNA fragment inserted and its orientation in the plasmid. Rapid, preparative isolation of the recombinant plasmid DNA was accomplished by using a modified Birnboim and Doly method (11), followed by centrifugation in a discontinuous, two-step cesium chloride-ethidium bromide gradient devised by Garger et al. (21). By using this protocol, covalently closed circular plasmid DNA can be separated from contaminating protein, RNA, and chromosomal DNA in 3 to 5 h. The resulting plasmid DNA is considerably freer of RNA than it is when purified by a 45-h equilibrium run in a homogeneous gradient, and it can be used directly for end labeling and DNA sequence analysis. Briefly, for the quick-seal centrifuge tube (5/8 by 3 in. [1.59 by 7.62 cm]; Beckman Instruments, Inc.), nucleic acid was dissolved to a final volume of 2.4 ml in TE (10 mM Tris-hydrochloride [pH 8.0], 1 mM EDTA), and 4.2 g of cesium chloride was added to this to obtain a weight of 6.6 g. Of a 10-mg/ml ethidium bromide solution, 0.4 ml was added to bring the final weight to 7.0 g and the cesium chloride to 60% (wt/wt). This nucleic acid-cesium chloride mixture was then layered beneath 8 ml of a 43% (wt/wt) cesium chloride solution. The tube was then filled to the top with 43% cesium chloride in TE. The sealed tube was run at 65,000 rpm for 5 h in a Beckman type 70.1 Ti rotor or at 70,000 rpm for 3 h in a Beckman type 80 Ti rotor at 20°C. Plasmid DNA was collected, ethidium bromide was removed, and DNA was precipitated as described previously (37). Purified plasmid DNA was subjected to cleavage by a large variety of commercially obtained restriction endonucleases by the specifications of the manufacturers. The resulting restriction map information was used in the Maxam and Gilbert approach to DNA sequencing described below.

DNA sequence analysis. DNA sequence analysis was performed by a modification of the chemical method of Maxam and Gilbert (37–39), including protocols for the removal of 5' phosphates with bacterial alkaline phosphatase and the labeling of 5' ends with [γ -³²P]ATP (1,000 to 3,000 Ci/mmol; New England Nuclear Corp.) and T4 polynucleotide kinase. Both plus and minus strands were sequenced to prevent errors in sequence determination caused by methylation or secondary structural features of the DNA. Polyacrylamide (6 or 8%)–urea (7 M) gels, 60 to 80 cm in length and 0.4 mm in

thickness, were used for resolution of sequence reactions (37). Sequence homologies between tubulin genes were compared by using the COMPAR program prepared by H. Selick and V. Bedian, University of Pennsylvania.

Primer extension experiments. DNA primers ranging in size from 25 to 150 bp were isolated from 5' regions of the tubulin genes by cutting with appropriate restriction enzymes, separating fragments on 6% polyacrylamide gels, and electroeluting the desired fragments. The 5' termini of these double-stranded fragments were labeled with [γ -³²P]ATP, using T4 polynucleotide kinase (as described above). After repeated lyophilizations to remove salt, the double-stranded primer was denatured at 90°C for 15 min in a formamide-containing buffer as described by Ghosh et al. (22). The RNA was heated for 1 to 2 min at 55°C, and the denatured DNA primer was added to the RNA. The RNA-DNA mixture was rapidly transferred to 52°C, and the annealing reaction was carried out for 12 to 16 h. The annealed mRNA-DNA primer was precipitated in 3 M ammonium acetate–70% ethanol, washed with 95% ethanol, and dried. With the exceptions that 0.5 mM replaced the 1 mM concentration of dNTP in the reaction mix and that actinomycin D was added at 60 μ g/ml, the extension reaction buffer was as described previously (22). Reverse transcriptase (Bethesda Research Laboratories) generally was used at 50 U per reaction for 2 to 8 μ g of polyadenylated [poly(A)⁺] RNA and 100 ng of labeled primer. Extended primers were run on sequencing gels, using sequenced DNA fragments for exact size markers. The gels were exposed directly to Kodak XAR-5 film, using DuPont Cronex intensifying screens at –70°C. A 1.5-bp correction was taken into account when comparing fragments which had not undergone chemical cleavage with those from DNA sequencing reactions.

Isolation of RNA. RNA was isolated from a variety of stages in the life cycle and cell cycle of *C. reinhardi* 137 C, including nondividing gametic cells, gametic cells induced to produce tubulin mRNA, asynchronous vegetative cells, and synchronous vegetative cells isolated during cell division. Vegetative cultures of *C. reinhardi* were grown either synchronously in a 12-h light-dark cycle or asynchronously in constant light. Gametes were prepared from vegetative cells by changing to a nitrogen-free media (61). Synchronous deflagellation of gametic cells by pH shock (60) was used to induce cells to the high levels of tubulin synthesis and tubulin mRNA production which accompany flagellar regeneration (44, 54, 60). Total RNA and poly(A)⁺ RNA were prepared as previously described (12).

RESULTS

Coding regions. In previous studies, using cDNA clones from tubulin mRNAs to probe Southern blots of both genomic DNA and DNA from clones of the four *C. reinhardi* tubulin genes (12), we mapped the approximate location of the coding regions for each gene. Using this information, we have constructed appropriate subclones of the 5' portions of the individual genes by insertion of DNA restriction fragments from the genomic clones into the plasmid vector pBR322 or pUC9. After extensive restriction enzyme mapping of these subclones, the DNA within and flanking the 5' region was sequenced by the method of Maxam and Gilbert (37–39). Figure 1 displays the regions flanking the ATG translation initiator codon in the genes which we previously designated as the α_1 -, α_2 -, β_1 -, and β_2 -tubulin genes (12).

In *C. reinhardi*, the α -tubulin genes do not cross-hybridize with the β -tubulin genes (12). This lack of homology is

α_1 ⁺⁵⁰ CCCAGTTGGGACCGCGGTCATTACCCCTCGCGAACCCAGCGTAGGA⁺¹⁰⁰AAAGCTCCCTTCTCTCCCTGATTCTTCTCCGTCTCCCTTGGCAAC⁺¹⁵⁰CCGTTGGTACGCGAGGTCATTTAAGCCCTCGCGTCACCCCGGTTAGCAGAGCTCCCCCTTTCGCTCCCTTCTCTCTCCGTATCCAGTCCCTTGGCAAC⁺⁵⁰
 α_2 ⁺¹⁵⁰ ATG CGT GAG GTC ATC TCC ATC CAC ATC ATC GGC CAG GCC GGT ATC CAG GTC...intron
Met Arg Glu Val Ile Ser Ile His Ile Gly Gln Ala Gly Ile Gln
ATG CGT GAG GTC ATC TCC ATC CAC ATC ATC GGC CAG GCC GGT ATC CAG GTC...intron
⁺¹⁵⁰

B1 CGTGAAGGCCTTGACCAAAĠCAGCGAGACTTTCGCTĠCCCGCGTCCĠĠCCCCCĠCTCTTGGGCTCTCGCTTTCĠĠGGCGAGCCĠĠĠAGCAAAAC
 +100
 B2 CCTAGATCACTACCACCTTCTACTACAGĠCCACTCGAGĠTTGTGATCGĠACTCCGCTAĠAGGGGGCCCTCTTCTCTTCTTCACTĠĠCAACCĠGCAĠAC
 +50
 B1 ĠATG CGT GAG AĠC GTC CAC ATĠ CAG GTG...intron...CAG ĠGT GGC CAG TĠC GGC AAC CAĠ ATT GGT GCT ĠAG TTC
 Met Arg Glu Ilu Val His Ilu Glu Gly Gly Glu Cys Gly Asn Glu Ilu Gly Ala Lys Phe
 B2 ATG CGT GĠG ATC GTC CAĠ ATC CAG GTG...intron...CAG GGT GGC CAG TGC GGC AAĠ CAG ATC GGT ĠCC AAG TTC
 +150
 B1 +350 TGG GAG GTC GTC TCC ĠAC GAG CAG CAC ĠGC ATC GAC ĠAC ACC GGC ACC ACĠ TAC CAT GGC ĠAC TCC GAC CTĠ CAG...
 Trp Glu Val Val Ser Asp Glu His Gly Ilu Asp Pro Thr Gly Ser Tyr His Gly Asp Ser Asp Leu Glu
 B2 TĠG GAG GTC GTĠ TCC GAC GAG ĠAC GGC ATC ĠAC ACC GGC ACC TAC CAT ĠC GGC GAC TCA ĠAC CTG CAG...
 +350 +400 +400

FIG. 1. Portions of the amino acid coding and noncoding regions flanking the ATG and β -tubulin genes of *C. reinhardtii*. The sequences shown were obtained by using a modified Maxam and Gilbert protocol (37–39). Position numbers are relative to the mRNA initiation sites. Arrows indicate regions of homology between α - and β -tubulin genes or between β - and β -tubulin genes. The brackets enclose a CT-rich region present in all four genes. Notice that, in the comparison of β - with β -tubulin, single bp changes in the coding region have not altered the codon specificity at that position. Intron sequences have been omitted from the figure except for the intron-exon junctions. The sequences of the first introns present in the β - and β -tubulin genes are as follows.

IVS₁_{B1} GTGGCGCAAGCTTCGCGTCTGGCTTAGGCTGCTGGGGGAAAGAGCCCGCGCTCGCAAGATCTTCGGAGTTCACATCGTCTCGGCGTCACTCAATTTGCGCTTCTCGTCCCAATTGCAATCTCAG

IVS₁_{B2} GTGGCGTTGAGCGCTTAGCGCATTTGGCTGAGGCTAGCGCAGTCAAGGGGCGCGGGCTCGTGCGTACACCCCCCGGGCTCAATTTCAAACCTGTTCCGACTTCGAGGCTCATCGTCGCTCGGCTTGCGCCTTACATTACAG

clearly evident in a comparison of the nucleotide sequences of the two gene sets presented in Fig. 1 and 2. No stretch of nucleotide homology is longer than 10 bp for the transcribed regions given here. However, comparisons of α_1 - and α_2 -tubulin genes or β_1 - and β_2 -tubulin genes reveal marked conservation of nucleotide sequences in the coding regions (Fig. 1). The β_1 - β_2 gene set differs in sequence by only 3 bp in the first 43 amino acid codons (129 bp), and all of these differences are silent with respect to amino acid coding.

The α_1 - and α_2 -tubulin genes contain an intron between amino acid codons 15 and 16, whereas the β_1 - and β_2 -tubulin genes contain an intron between the amino acid codons 8 and 9. The sequences in the intron-exon border regions are strongly conserved. As can be seen for the first intron in the β_1 - and β_2 -tubulin genes (legend to Fig. 1), this sequence homology continues for only a short distance. The internal intron sequences of the β -tubulin genes bear little homology and differ in length by 20 bp. It should also be noted that intron placement differs between α - and β -tubulin genes (Fig. 1), and no homology is observed in their intron sequences with the exception of the conserved intron-exon borders.

In the noncoding region upstream of the ATG initiator codon, intermittent stretches of sequence homology can be noted within the α -tubulin gene set and to a lesser degree within the β -tubulin gene set (Fig. 1). Perhaps more striking, however, are two other features of this region. The first is the highly CT-rich nature of the mRNA sense strand of all four genes in a stretch extending from ca. 50 to 5 bp upstream from the ATG codon. In the case of the α_1 -tubulin gene, 39 of 41 nucleotides in the positive strand in this region are C or T. The second feature is the presence of a highly conserved consensus hexanucleotide, GCAA(A/C)C, immediately preceding the initiator codon. What influence either or both of these features have on tubulin mRNA structure, ribosome binding, and translational efficiency is unknown. It is interesting to note, however, that in the identical position occupied by the *C. reinhardi* GCAA(A/C)C sequence, the rat α -tubulin gene contains the sequence GCAACC (35) and the human β -tubulin gene has the related sequence TTAACC (33).

Promoter regions. Examination of sequences upstream of the ATG initiator codon in the four tubulin genes reveals a number of potential TATA box sequences that could be used for aligning RNA polymerase II molecules with the transcription initiation sites (Fig. 2). Therefore, to map the cap site for each gene, both primer extension (22) and S1 nuclease digestion (9, 59) experiments were carried out. The primer extension strategy used for locating the β_2 -tubulin mRNA initiation site is outlined in the diagram accompanying Fig. 3. Using a similar strategy for all primer extension experiments, we isolated and labeled short restriction enzyme fragments within the noncoding region for each gene, annealed the radioactively labeled primer to mRNA, and extended each primer to the 5' terminus of its respective mRNA by using avian myeloblastosis virus reverse transcriptase. Primer extension data for the β_2 -tubulin gene (Fig. 3) show that the primary transcription initiation site for this gene is at a deoxyadenosine residue located 132 bp upstream from the ATG initiator codon and 30 bp downstream from a 7- or 8-bp-long TATA box sequence. Primer extension results for the α_1 -tubulin gene (data not shown) position the major mRNA cap site for this gene 140 bp upstream from the ATG initiator codon.

As will be discussed in greater detail below, the β_1 -tubulin gene contains two regions in close proximity, each of which

has the major characteristics we have found for tubulin gene promoters. The primer extension data (Fig. 4a) show that the deoxyadenosine residue 159 bp from the ATG initiator codon is the preferred cap site for the β_1 -tubulin gene and that the upstream promoter region is the one used exclusively for mRNA production in the nondividing gametic cells of *C. reinhardi* used here.

The α_2 -tubulin gene also has two potential promoter regions (Fig. 2). One such region is located far upstream (541 bp from the initiator ATG), whereas the location of the other possible α_2 -tubulin promoter puts the mRNA cap site 141 bp upstream from the initiator ATG codon, a position more typical of the α_1 -, β_1 -, and β_2 -tubulin genes. Initially, we questioned the potential effectiveness of the downstream promoter because of the presence of a deoxycytidine residue in its short TATA box sequence. Results of primer extension studies (Fig. 4b) and S1 nuclease protection experiments (data not shown) favored the downstream promoter as the site of transcription initiation. Nevertheless, neither experiment could rule out the far-upstream promoter of the α_2 -tubulin gene as the true promoter region should the unprocessed transcript contain an intervening sequence in its 5' nontranslated region. We therefore prepared a synthetic oligonucleotide primer complementary to a unique 5' region of the α_2 -tubulin mRNA (position +79 to +94, Fig. 1), end labeled the primer with [γ - 32 P]ATP, and used reverse transcriptase to produce a short single-stranded cDNA that could be sequenced. The resulting sequence data (Fig. 4c) show that the downstream promoter region is the exclusive start site for α_2 -tubulin mRNA transcription in nondividing gametic cells. More recent experiments with RNA isolated both from synchronous and asynchronous vegetative cultures have yielded essentially identical primer extension data to that presented in Fig. 4b and c. These experiments demonstrate that the α_2 -pseudopromoter is also not used to any measurable extent for tubulin mRNA synthesis during cell division or during any other portion of the *C. reinhardi* cell cycle.

Potential regulatory sequences. To determine whether we could detect DNA sequences other than the TATA box sequences that might play a role in the transcription and coordinate control of the four *C. reinhardi* tubulin genes, we sequenced far upstream from the coding region of each gene (e.g., greater than 1,000 bp upstream in β_1 and in α_2). Computer analysis revealed several elements of potential interest, such as direct and inverted repeats, palindromic sequences, and dyads of symmetry, a number of which are noted in Fig. 2. However, most of these elements are present in only one or two of the four genes and are therefore less likely to be important in the regulation of the entire complement of tubulin genes. Nevertheless, two distinguishing features in the region 5' to the transcription initiation site deserve special note. The first is a 10- to 11-bp-long sequence between the TATA box and cap sites that is composed almost exclusively of G and C residues and is anchored at either end by deoxyadenosine residues. The second distinguishing element in the 5' promoter region is a 16-bp consensus sequence, GCTC(G/C)AAGGC(G/T)(G/C) (C/A)(C/A)G. Each of the four tubulin genes contains multiple copies of this consensus sequence within 150 bp upstream of the transcription start site (Fig. 2 and Table 1). In all cases, one or more copies in each gene bear 73% or greater homology to the consensus sequence. The consensus sequence itself was derived from the 15 16-bp sequences which exhibit the strongest homology (at least 67%) in our analysis (Table 1).

FIG. 2. 5' flanking sequences for the four tubulin genes and for a pseudopromoter region in the α_2 -tubulin gene. Position numbers are relative to the mRNA initiation site (+1) which are known sites for α_1 -, α_2 -, β_1 -, and β_2 -tubulin genes (Fig. 3 and 4). Important sequence elements are indicated as follows: TATA boxes, solid brackets; consensus sequence elements with high homology, solid boxes; consensus sequence elements with weak homology, dashed boxes; GC-rich regions, underscored by + signs; centers of palindromic sequences, asterisks; centers of dyads of symmetry, vertical arrows; limits of dyads of symmetry, dotted vertical lines; repeated sequences, horizontal arrows; absence of a C/G residue in β_1 -tubulin pseudopromoter, caret. The cap site indicated for the α_2 -tubulin pseudopromoter region (bottom line) is included to align the promoter region sequences and is not known to correspond to a mRNA initiation site. Its position relative to the authentic α_2 -tubulin gene cap site is -400. Likewise, a cap site for a pseudopromoter for the β_1 -tubulin gene (at position +58) is indicated by a cross above the potential cap site, but again there is no evidence for a mRNA initiating at this position.

FIG. 3. Determination of the transcription initiation site for β_2 -tubulin mRNA by primer extension analysis. The diagram outlines the step-by-step protocol used in primer extension experiments. In the case of the β_2 -tubulin gene, a fragment of 30 bp extending from the *NarI* to *XhoI* restriction enzyme sites was isolated from a plasmid containing the β_2 -tubulin gene. The 5' termini of this double-stranded fragment was labeled with [γ - 32 P]ATP. After heat denaturation at 90°C for 15 min, the anti-sense strand was annealed to tubulin mRNA. By using reverse transcriptase and unlabeled deoxynucleotides, the annealed primer was extended to the 5' end of the mRNA. The radioactive extension product was then isolated and run on an 8% polyacrylamide-7 M urea gel for sizing. For precise nucleotide markers, a *NarI* restriction fragment extending through the region of the cap site was isolated, end labeled at the 5' overhang of the same *NarI* site as the primer, cut with *BstNI* to produce a 138-bp fragment, and then subjected to DNA sequence reactions. A 1.5-bp correction was taken into account when comparing sequenced marker bands with primer extension bands. Comparison of the length of the extended primer with the sequenced *BstNI*-*NarI* restriction fragment allowed assignment of the primary transcription initiation site designated as +1 in the DNA sequence shown on the left. Lanes 1, 2, 3, and 4, Sequence reaction products of the 138-bp *BstNI*-*NarI* fragment from the anti-sense strand of the β_2 -tubulin gene (A+G, G, C, C+T, respectively). The sequence of the anti-sense strand and the corresponding sense strand are shown to the left of the gel. Lane 5, Extension with 30-bp primer (*XhoI*-*NarI* fragment [position +68 to +98, Fig. 1] of the β_2 -tubulin gene 32 P-labeled at the same 5' *NarI* overhang as in the marker lanes), using poly(A)⁺ mRNA from uninduced cells. Lane 6, Extension with the 30-bp end-labeled primer, using poly(A)⁺ mRNA from induced cells. Lane 7, Extension with 30-bp primer without mRNA (control lane). Note that even when several-fold greater amounts of uninduced cell mRNA (lane 5) are used in comparison with induced cell mRNA (lane 6), we still do not detect the additional band seen in lane 6.

DISCUSSION

In the present studies of the four *C. reinhardtii* tubulin genes, α_1 , α_2 , β_1 , and β_2 , we have identified the amino acid coding regions, the transcription initiation sites, potential regulatory sequences, and, in two genes, pseudopromoter sequences near the 5' termini of the genes.

From DNA sequencing of the tubulin gene coding regions we have been able to show that there is a strong conservation in amino acid sequence between the α - and β -tubulins of *C. reinhardtii* and the amino acid sequence of their counterparts in chickens (58), humans (33), rats (35), sea urchins (36), pigs (30), and yeasts (46). Only 1 amino acid residue of the first 15 in *C. reinhardtii* β -tubulin differs from that in β -tubulin from chickens (58), whereas 3 of 15 residues differ in the α -tubulin subunit. We have previously noted (12, 44) that

conservation of nucleotide sequence between the coding regions of the chicken α - (or β -) tubulin gene and the corresponding regions in the *C. reinhardtii* α - (or β -) tubulin gene permits strong cross hybridization with heterologous probes. Although both amino acid codon specificity and nucleotide sequence show a high degree of conservation between species for the α - or β -tubulin subunit, the placement of introns differs in the 5' portion of the *C. reinhardtii* α - or β -tubulin gene in comparison with intron placement in the rat α -tubulin (35) or human β -tubulin (33) gene, respectively. This difference is of interest in light of recent observations that introns may separate the coding regions for distinct polypeptide domains and may correlate with amino acid residues at the protein surface (15, 26).

The regions just upstream of the transcription initiation sites of the *C. reinhardtii* tubulin genes contain two highly

FIG. 4. Initiation site determination for the β_1 - and α_2 -tubulin mRNAs by primer extension experiments and DNA sequence analysis. Panels a and b show primer extension experiments similar to those described in the legend to Fig. 3. In these cases, the size of the extended primers has been determined by comparison with the sequence reaction products from restriction fragments of known DNA sequence. The numbers to the side of each gel designate the nucleotide positions from the ^{32}P -labeled end of each restriction fragment. These numbers do not include the 1.5-bp correction which must be taken into account when comparing sequence reaction products with the unreacted DNA fragments of the extended primers. The 80-cm polyacrylamide gels used in the primer extension experiments were cut into a top (T) and bottom (B) section for autoradiography. (a) Primer extension for β_1 -tubulin gene. Lanes 1, 2, 3, and 4, Sequence reaction products of the 781-bp *HindIII-PstI* restriction fragment from plasmid pBR322, which was ^{32}P -labeled at the 5' overhang of the *HindIII* site (G, A+G, C, C+T, respectively). Lane 5, Primer extension with the absence of mRNA in the extension reaction (control) and using a 35-bp primer ^{32}P -labeled at 5' termini (*StuI-Thal* fragment, position +68 to +103 in the β_1 -tubulin gene, Fig. 1). Lane 4, Primer extension with the 35-bp primer and poly(A)⁺ mRNA from induced cells. Lane 5, Primer extension with the 35-bp primer and total RNA from induced cells. Lane 5, 6, and 7 were purposely overexposed to demonstrate that, other than primer artifacts present in control and experimental lanes at levels which reflect the concentration of primer, all primer extension bands correspond to an mRNA initiation site in the β_1 upstream promoter region (Fig. 2). The arrow on the right indicates the expected position of the mRNA initiation site if the downstream β_1 -tubulin pseudopromoter were used. No clearly distinguishable band of radioactivity corresponding to the predicted transcript from the pseudopromoter can be detected in this gel or in other gels using different primer fragments. (b) Primer extension for the α_2 -tubulin gene. Lane 1, Control extension without mRNA, using a 30-bp 5' ^{32}P -end-labeled primer prepared from the α_2 -tubulin gene (*RsaI-NciI*, position +52 to +82, Fig. 1). Lane 2, Primer extension with the same primer as that in lane 1, using poly(A)⁺ mRNA from deflagellated gametic cells in the extension reaction. Lanes 3, 4, 5, and 6, Sequenced markers for size determination of extended primer (A+G, G, C, C+T, respectively). In this case, the plasmid used for markers is a pUC9 derivative containing a 269-bp *Thal* restriction cut of the α_2 -tubulin gene with an 8-bp *Sall* linker sequence added to each end and inserted into the *Sall* site of pUC9. The constructed plasmid was cut and ^{32}P end labeled at the *HindIII* site in the pUC9 polylinker region. A 317-bp *HindIII-EcoRI* fragment was isolated and used in sequence reactions. The upstream pseudopromoter for the α_2 -tubulin gene does not appear to be utilized as can be seen by the total absence of extension bands in the upper section of the gel. (c) DNA sequence of extended primer for α_2 -tubulin. A single-stranded 16-bp oligomer specific for the α_2 -tubulin gene (position +79 to +94, Fig. 1) was end labeled with [γ - ^{32}P]ATP and used as the primer in a primer extension with poly(A)⁺ mRNA from induced cells. This single-stranded primer was added directly to the extension reaction mixture without the denaturation and annealing reactions required for a double-stranded primer. The resulting labeled extension product was separated from free primer by gel electrophoresis and then removed from the gel by electroelution. The extended primer was sequenced by the Maxam and Gilbert chemical method (37-39). Sequences for the 3' end of the anti-sense strand and the corresponding 5' end of the sense strand are indicated to the right of the gel. (Lanes, 1, 2, 3, and 4 correspond to A+G, G, C, C+T sequencing reactions, respectively.)

conserved features (in addition to the usual TATA box), which potentially may function in the regulation of transcription. The first is a highly GC-rich area of 10 to 11 bp in length located between the TATA box and the cap site in each gene

(Fig. 2). A number of eucaryotic genes, including some encoding tubulin and other structural proteins, have been shown to contain GC-rich sequences in this region (7, 20, 35). The exceptional length of the GC stretch in the case of

TABLE 1. Consensus sequence for the four tubulin genes^a

Gene position	Nucleotide sequence	% Homology	
		10 bp	16 bp
α_1	-55 G C C C C C A T T C C G C G C G C G	60	67
	-76 G C T T C G A A G T C C G C G A T C	90	80
	-84 G C T C C G A A G C C T C G A A G	100	100
α_2	-46 G C C C C G A T G C A C T A C T	70	67
	-63 G C G C G A T G G T C C T T A G	70	67
	-78 G C T C C A T G C G C C C G C G	70	73
	-147 T T T G C A A G C C A T G C A G	70	67
β_1	-65 G C T C C A A G C A G C T C G T	100	80
	-73 G C T C C A A T G C T C G A A G	80	87
	-81 C T T C G A A T G C T G C A A T	70	73
β_2	-48 G C T C C A G G C G C A G C G C	90	73
	-58 G C T C C G T G C G C C T C C A	80	73
	-77 G C T T C G G G C T G C A T G	60	67
	-87 G C C C G A A G C C C T T C G	70	67
	-127 G C T T C C G C G C G C T G C A	70	67
G 11 1 3 6 3 2 11 11 1 6 5 5 3 2 8			
C 3 13 3 10 9 1 1 3 13 2 8 4 4 6 2			
A 11 8 3 1 6 5 2			
T 1 2 11 2 4 4 1 1 4 1 6 2 2 3			
Consensus G C T C C G A A G C C G - C C G			
Consensus sequence found in pseudopromoter regions			
α_2	-57 G C T C C G A T G T G A G G T C C	70	60
	-117 G T T C G A T G A C A T T C C G	70	67
	β_1 -48 C C T C C (*) A A G C A C A T A C T	60 (50)	60 (47)

^a In analyzing the regions upstream of the mRNA initiation site (+1) for the four genes, the rules outlined by Davidson et al. (17) were used to determine the consensus sequence element. A number of nucleotide stretches were found which shared at least 60% homology to each other. Of these, the 15 showing the strongest homology (at least 67% to the 16-bp stretch) were used to determine the consensus sequence. The tabulation below the sequences summarizes the nucleotide content of these 15 sequences. In those cases in which a nucleotide pair has been assigned to a particular position in the consensus sequence, that assignment is based on a 67% or greater combined incidence of these two nucleotides at that position. A hyphen has been used in position 13 of the consensus sequence since the only distinguishing feature of this position is the absence of an A residue. The homology of each repeated sequence element to a 10-bp core (the left side of the consensus sequence) and to the entire 16-bp stretch is given in the columns on the right. At the bottom of the table are listed three nucleotide stretches showing some homology to the consensus sequence found in pseudopromoter regions of the α_2 - and β_1 -tubulin genes. Their positions are relative to potential but unused cap sites. In the case of the β_1 -pseudopromoter, the position is given relative to a potential cap site located at +58 in the 5' noncoding region of the β_1 -tubulin gene (Fig. 2). The β_1 -tubulin pseudopromoter consensus sequence shows 60% homology to the consensus sequence if a deletion is assumed in the fifth position (indicated by an asterisk). Otherwise, the homology drops to 47%.

the *C. reinhardi* tubulin gene raises the prospect that this sequence may influence the efficiency of tubulin gene transcription either through its influence on promoter region structure (7) or by its ability to direct the binding of molecules which stimulate transcription or both.

The second distinguishing feature, and the one which we consider the most likely to influence coordinate regulation of the tubulin genes, is the 16-bp consensus sequence, GCTC(G/C)AAGGC(G/T)(G/C)–(C/A)(C/A)G (Fig. 2 and Table 1). This sequence is found in three or more copies immediately upstream of the TATA box in each of the tubulin genes. It should be noted that the longer 16-bp sequence contains a shorter and more highly conserved 10-bp "core" sequence [GCTC(G/C)AAGGC]. Judging from

the short consensus sequences found to be important in the regulation of other eucaryotic genes (10, 17, 18, 50), the 10-bp core sequence of the *C. reinhardi* tubulin genes may be of adequate length to serve as a regulatory element.

The *C. reinhardi* consensus sequence elements lie just upstream from the TATA box in a position similar to that occupied by putative or proven regulatory elements in other eucaryotic genes (10, 18, 25, 28, 29, 41, 45, 49, 50, 55, 56). The furthest upstream element we have detected is at –147 bp. A striking feature in all of the tubulin genes is the tight clustering and frequent overlapping of consensus sequences in the –90 to –35 regions (Fig. 2). No sequences with even moderate homology to the 10- or 16-bp consensus sequence can be found in the promoter regions of sequenced tubulin genes from other organisms (33, 35).

In the α_1 - and β_2 -tubulin genes, we find the consensus elements in association with sequences which contain dyads of symmetry, a feature also found in the promoter region of many of the heat shock protein genes of *Drosophila* strains (49). Another feature common to both α_1 - and β_2 -tubulin genes but not present in the other two tubulin genes is an area of close homology in the cap site region (GCCCGGATTGCAAGACATT). A possible role for these shared features in the selective transcription of this α_1 - β_2 -tubulin gene set during the cell cycle or developmental life cycle of *C. reinhardi* is presently under investigation.

Two tubulin genes, β_1 and α_2 , contain regions that appear to be pseudopromoters since these regions are deceptively similar to the authentic tubulin gene promoters but do not produce detectable transcripts in response to tubulin gene induction.

Examination of the β_1 -tubulin gene pseudopromoter, which is located just downstream of the transcription initiation site of the gene (Fig. 2), shows that it contains a strong TATA box and a GC-rich sequence with spacing and orientation similar to that found in the authentic tubulin gene promoters. The only striking feature which discriminates this promoter region from the verified tubulin gene promoters is the absence of a sequence with strong homology to the 10- and 16-bp consensus sequences. Of interest, however, is a 15-bp sequence beginning 21 bp upstream from the TATA box of the β_1 -pseudopromoter which would conform to the consensus sequence if it contained a G or C residue at the fifth position in the sequence (Table 1). Since this is one of the more highly conserved positions in the 10-bp core consensus sequence, one contributing factor to the lack of function of the β_1 -tubulin pseudopromoter could be the absence of this single G/C base pair near the center of the consensus sequence.

The character of the pseudopromoter for the α_2 -tubulin gene is somewhat different (Fig. 2 [α_2]). It is located 400 bp upstream of the authentic promoter region in the α_2 -tubulin gene and contains all of the elements of the tubulin gene promoter regions that we have detailed above. Indeed, the TATA box sequence of the upstream α_2 -tubulin pseudopromoter is a much closer match to the usual TATA box consensus sequence (6) than is its downstream partner which contains a deoxycytidine residue, a highly unusual feature which is also shared by the TATA box sequence of the authentic β_1 -tubulin promoter. Alternative 5' promoter sites in several eucaryotic genes (8, 13, 27, 32, 62) at first suggested to us the possibility that the upstream and downstream promoter regions in the α_2 -tubulin gene might be an additional example of this phenomenon. However, as discussed above, the primer extension and sequencing data of Fig. 4b and c demonstrate that the α_2 -tubulin mRNA is

produced from the downstream promoter. The mRNA clearly is not synthesized from an upstream promoter with the subsequent removal of an intervening sequence in the 5' noncoding region, as is sometimes observed in other genes (47, 63). There is a marked conservation of specific nucleotides in particular positions of the consensus sequence (Table 1). Thus, one possible explanation for the lack of utilization of the α_2 -upstream promoter region is the substitution of alternative nucleotides for the G residue at position 9 and for the C residues at positions 2 and 10 of the consensus sequence elements of the α_2 -pseudopromoter. If the consensus sequence serves as a recognition site for regulatory molecules important for tubulin gene transcription, the repeats in the α_2 -pseudopromoter region may lack the requisite nucleotide pattern for productive interaction with these regulatory molecules. A complete understanding of why the authentic promoter regions of the tubulin genes are used, whereas the α_2 - and β_1 -pseudopromoters are not, will depend on future experiments in which the promoter regions are modified in vitro and tested for activity in vivo in transformed *C. reinhardi* cells.

If the search for regulatory elements is not confined to the regions upstream of the cap site, then one additional difference between the α_2 - and β_1 -tubulin pseudopromoters and the authentic tubulin promoters is the absence of a "G-deficient" region beginning near the transcription start site and extending downstream 40 to 60 bp. In view of the presence of regulatory elements within the transcribed region of certain genes (1, 31) and enhancer sequences in the intervening sequences of immunoglobulin genes (4, 23, 51), we presently cannot rule out the potential regulatory importance of the G-deficient regions, the short sequence homologies flanking the start of the coding region of the four tubulin genes, or even sequences within the intervening sequences of the genes. Likewise, we do not know whether the same or different regulatory sequences are involved in the production of tubulin during cell division in synchronized cultures of *C. reinhardi* (2, 12, 61). The regulation of cell cycle-dependent histone synthesis in yeasts by apparent autonomously replicating sequence elements that are located 3' to certain of the histone genes (48) suggests the possibility that the control of tubulin production during *C. reinhardi* cell division could also rely on sequences that lie outside the regions we have currently analyzed. These possible exceptions notwithstanding, the preponderance of data accumulated to date on the regulation of eucaryotic genes points clearly to the region immediately upstream of the transcription start site as playing a predominant role in the control of gene transcription (3, 6, 16, 40, 42, 57). In a number of systems (10, 18, 25, 49, 50, 56) which, like the tubulin gene system in *C. reinhardi*, show rapid and coordinate induction of mRNA synthesis, short consensus sequences just upstream of the TATA box have been identified and verified as the key elements regulating gene transcription. The presence of an analogous 16-bp consensus sequence in a similar position in all four of the *C. reinhardi* tubulin genes suggests a potential role for this sequence in controlling the coordinate expression of the spatially separated members of this gene set. As with other eucaryotic gene systems (10, 18, 25, 49, 50, 56), the ultimate definition of the sequences important to the coordinate regulation of the *C. reinhardi* tubulin genes will depend on the analysis of genes modified in specific sequences and reintroduced into cells by genetic transformation.

ACKNOWLEDGMENTS

We thank A. Marcus, R. Perry, and S. Tilghman for the critical reading of this manuscript; L. Gaitmaitan for teaching us DNA

sequencing; R. Scott for help with primer extension protocols; and N. Malin, P. Dietrich, S. Hazel, and S. Chow for technical assistance. We also thank Jane Hutchens, Cathy Cannon, and Fred Kalish for help in the preparation of the manuscript.

K.J.B. was supported by a postdoctoral training grant from the National Institutes of Health. This work was supported by the National Science Foundation (grant PCM-8302639), Zoecon Corp., and a grant from the Commonwealth of Pennsylvania.

LITERATURE CITED

1. Anziano, P. Q., D. K. Hanson, H. R. Mahler, and P. S. Perlman. 1982. Functional domains in introns: *trans*-acting and *cis*-acting regions of intron 4 of the cob gene. *Cell* 30:925-932.
2. Ares, M., Jr., and S. H. Howell. 1982. Cell cycle stage-specific accumulation of mRNAs encoding tubulin and other polypeptides in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U.S.A.* 79:5577-5581.
3. Baker, C., and E. J. Ziff. 1981. Promoters and heterogeneous 5' termini of the messenger RNAs of adenovirus serotype 2. *J. Mol. Biol.* 149:189-221.
4. Banerji, J., L. Olson, and W. Schaffner. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33:729-740.
5. Barta, A., R. I. Richards, J. D. Baxter, and J. Shine. 1981. Primary structure and evolution of rat growth hormone gene. *Proc. Natl. Acad. Sci. U.S.A.* 78:4867-4871.
6. Benoist, C., and P. Chambon. 1981. In vivo sequence requirements of the SV40 early promoter region. *Nature (London)* 290:304-310.
7. Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss. 1983. Physical characteristics in eucaryotic promoters. *Nucleic Acids Res.* 11:4521-4540.
8. Benyajati, C., N. Spoerel, H. Haymerle, and M. Ashburner. 1983. The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5' end in different developmental stages. *Cell* 33:125-133.
9. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
10. Bienz, M., and H. R. B. Pelham. 1982. Expression of a *Drosophila* heat-shock protein in *Xenopus* oocytes: conserved and divergent regulatory signals. *EMBO J.* 1:1583-1588.
11. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.
12. Brunke, K. J., E. E. Young, B. U. Buchbinder, and D. P. Weeks. 1982. Coordinate regulation of the four tubulin genes of *Chlamydomonas reinhardi*. *Nucleic Acids Res.* 10:1295-1310.
13. Carlson, M., and D. Botstein. 1982. Two differentially regulated mRNAs with different 5' ends encode secreted and intracellular forms of yeast invertase. *Cell* 28:145-154.
14. Cochet, M., A. C. Y. Chang, and S. N. Cohen. 1982. Characterization of the structural gene and putative 5'-regulatory sequences for human proopiomelanocortin. *Nature (London)* 297:335-338.
15. Craik, C. S., S. Sprang, R. Fletterick, and W. J. Rutter. 1982. Intron-exon splice junctions map at protein surfaces. *Nature (London)* 299:180-182.
16. Darnell, J. E., Jr. 1982. Variety in the level of gene control in eukaryotic cells. *Nature (London)* 297:365-371.
17. Davidson, E. H., H. T. Jacobs, and F. J. Britten. 1983. Very short repeats and coordinate induction of genes. *Nature (London)* 301:468-470.
18. Donahue, T. F., R. S. Daves, G. Lucchini, and G. R. Fink. 1983. A short nucleotide sequence required for regulation of HIS4 by the general control system of yeast. *Cell* 32:89-98.
19. Donehower, L. A., A. L. Huang, and G. L. Hager. 1981. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. *J. Virol.* 37:226-238.
20. Everett, R. D., D. Baty, and P. Chambon. 1983. The repeated

- GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res.* **11**:2447-2464.
21. Garger, S. J., O. M. Griffith, and L. K. Grill. 1983. Rapid purification of plasmid DNA by a single centrifugation in a two-step cesium chloride-ethidium bromide gradient. *Biochem. Biophys. Res. Commun.* **117**:835-842.
 22. Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman. 1978. Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. *J. Mol. Biol.* **126**:813-846.
 23. Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**:717-728.
 24. Grez, M., H. Land, K. Giesecke, G. Schutz, A. Jung, and A. E. Sippel. 1981. Multiple mRNAs are generated from the chicken lysozyme gene. *Cell* **25**:743-752.
 25. Hinnebusch, A. G., and G. R. Fink. 1983. Repeated DNA sequences upstream from HIS1 also occur at several other co-regulated genes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **258**:5238-5247.
 26. Inana, G., J. Piatigorsky, B. Norman, C. Slingsby, and T. Blundell. 1983. Gene and protein structure of a β -crystallin polypeptide in murine lens: relationship of exons and structural motifs. *Nature (London)* **302**:310-315.
 27. Jellinghaus, U., U. Schatzle, W. Schmid, and W. Roewekamp. 1982. Transcription of a *Dictyostelium* discoidin-I gene in yeast. *J. Mol. Biol.* **159**:623-636.
 28. Jones, C. W., and F. C. Kafatos. 1980. Structure, organization and evolution of developmentally regulated chorion genes in a silkworm. *Cell* **22**:855-867.
 29. Karin, M., and R. I. Richards. 1982. Human metallothionein genes—primary structure of the metallothionein-II gene and a related processed gene. *Nature (London)* **299**:797-802.
 30. Kraus, E., M. Little, T. Kempf, R. Hofer-Warbinek, W. Ade, and H. Ponstingl. 1981. Complete amino acid sequence of β -tubulin from porcine brain. *Proc. Natl. Acad. Sci. U.S.A.* **78**:4156-4160.
 31. Langford, C. J., and D. Gallwitz. 1983. Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. *Cell* **33**:519-527.
 32. Langridge, P., and G. Feix. 1983. A zein gene of maize is transcribed from two widely separated promoter regions. *Cell* **34**:1015-1022.
 33. Lee, M. G.-S., S. A. Lewis, C. D. Wilde, and N. J. Cowan. 1983. Evolutionary history of a multigene family: an expressed human β -tubulin gene and three processed pseudogenes. *Cell* **33**:477-487.
 34. Lefebvre, P. A., S. A. Nordstrom, J. E. Moulder, and J. L. Rosenbaum. 1978. Flagellar elongation and shortening in *Chlamydomonas*. *J. Cell Biol.* **78**:8-27.
 35. Lemischka, I., and P. A. Sharp. 1982. The sequences of an expressed rat α -tubulin gene and a pseudogene with an inserted repetitive element. *Nature (London)* **300**:330-335.
 36. Luduena, R. F., and D. O. Woodward. 1973. Isolation and partial characterization of α - and β -tubulin from outer doublets of sea-urchin sperm and microtubules of chick-embryo brain. *Proc. Natl. Acad. Sci. U.S.A.* **70**:3594-3598.
 37. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 38. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**:560-564.
 39. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
 40. McGinnis, W., A. W. Shermoen, and S. K. Beckendorf. 1983. A transposable element inserted just 5' to a *Drosophila* glue protein gene alters gene expression and chromatin structure. *Cell* **34**:75-84.
 41. McKnight, S., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* **217**:316-324.
 42. McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel. 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. *Cell* **25**:385-398.
 43. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 44. Minami, S. A., P. S. Collis, E. E. Young, and D. P. Weeks. 1981. Tubulin induction in *C. reinhardtii*: requirement for tubulin mRNA synthesis. *Cell* **24**:89-95.
 45. Muskavitch, M. A. T., and D. S. Hogness. 1982. An expandable gene that encodes a *Drosophila* glue protein is not expressed in variants lacking remote upstream sequences. *Cell* **29**:1041-1051.
 46. Neff, N. F., J. H. Thomas, P. Grisati, and D. Botstein. 1983. Isolation of the β -tubulin gene from yeast and demonstration of its essential function in vivo. *Cell* **33**:211-219.
 47. Ordahl, C. P., and T. A. Cooper. 1983. Strong homology in promoter and 3'-untranslated regions of chick and rat α -actin genes. *Nature (London)* **303**:348-349.
 48. Osley, M. A., and L. Hereford. 1982. Identification of a sequence responsible for periodic synthesis of yeast histone 2A mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **79**:7689-7693.
 49. Pelham, H. R. B. 1982. A regulatory upstream promoter element in the *Drosophila* hsp 70 heat-shock gene. *Cell* **30**:517-528.
 50. Pelham, H. R. B., and M. Bienz. 1982. A synthetic heat-shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene. *EMBO J.* **1**:1473-1477.
 51. Queen, C., and D. Baltimore. 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell* **33**:741-748.
 52. Remillard, S. P., and G. B. Witman. 1982. Synthesis, transport, and utilization of specific flagellar proteins during flagellar regeneration in *Chlamydomonas*. *J. Cell Biol.* **93**:615-631.
 53. Rochaix, J.-D., and J. van Dillewijn. 1982. Transformation of the green alga *Chlamydomonas reinhardtii* with yeast DNA. *Nature (London)* **296**:70-72.
 54. Silflow, C. D., and J. L. Rosenbaum. 1981. Multiple α - and β -tubulin genes in *Chlamydomonas* and regulation of tubulin mRNA levels after deflagellation. *Cell* **24**:81-88.
 55. Snyder, M., M. Hunkapiller, D. Yuen, D. Silver, J. Fristrom, and N. Davidson. 1982. Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* **29**:1027-1040.
 56. Struhl, K. 1982. Regulatory sites for his3 gene expression in yeast. *Nature (London)* **300**:284-287.
 57. Tsuda, M., and Y. Suzuki. 1981. Faithful transcription initiation of fibroin gene in a homologous cell-free system reveals an enhancing effect of 5' flanking sequence far upstream. *Cell* **27**:175-182.
 58. Valenzuela, P., M. Quiroga, J. Zaldivar, W. J. Rutter, M. W. Kirschner, and D. W. Cleveland. 1981. Nucleotide and corresponding amino acid sequences encoded by α - and β -tubulin mRNAs. *Nature (London)* **289**:650-659.
 59. Weaver, R. F., and C. Weissmann. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15 S β -globin mRNA precursor and mature 10 S β -globin mRNA have identical map coordinates. *Nucleic Acids Res.* **7**:1175-1193.
 60. Weeks, D. P., and P. S. Collis. 1976. Induction of microtubule protein synthesis in *Chlamydomonas reinhardtii* during flagellar regeneration. *Cell* **9**:15-27.
 61. Weeks, D. P., and P. S. Collis. 1979. Induction and synthesis of tubulin during the cell cycle and life cycle of *Chlamydomonas reinhardtii*. *Dev. Biol.* **69**:400-407.
 62. Young, R. A., O. Hagenbuehle, and U. Schibler. 1981. A single mouse α -amylase gene specifies two different tissue-specific mRNAs. *Cell* **23**:451-458.
 63. Zakut, R., M. Shani, D. Givol, S. Neuman, D. Yaffe, and U. Nudel. 1982. Nucleotide sequence of the rat skeletal muscle actin gene. *Nature (London)* **298**:857-859.