Vertebrate and Nematode Genes Coding for Yolk Proteins Are Derived from a Common Ancestor†

Denise Nardelli,‡ Susan Gerber-Huber,‡,§ Fred D. van het Schip,∥ Max Gruber,∥ Geert AB,∥ and Walter Wahli*,‡
Institut de Biologie animale, Université de Lausanne, Bâtiment de Biologie, CH-1015 Lausanne, Switzerland, and Biochemisch Laboratorium, University of Groningen, 9747 AG Groningen, The Netherlands
Received April 16, 1987; Revised Manuscript Received May 27, 1987

ABSTRACT: One of the most obvious characteristics of the egg cells of oviparous animals is their large size resulting to a major extent from the deposition of nutritional reserves, mainly constituted of yolk proteins. In general, these are derived from a precursor called vitellogenin, which undergoes posttranslational modifications during secretion and during transport into and storage within the oocytes. Comparative analysis of the structural organization of the vitellogenin gene and of its product in different species shows that the vitellogenin gene is very ancient and that in vertebrates the gene may have more resemblance to the earliest gene than in invertebrates.

Synthesis of vitellogenin, the precursor of yolk proteins, is tissue specific and strictly regulated by hormones, or in several invertebrate species by yet unidentified factors (Wallace, 1985; Wahli & Ryffel, 1985; Bownes, 1986). In insects, vitellogenin is produced under the control of juvenile hormone and ecdysone in the female fat body. In *Drosophila*, it is also produced in the follicle cells of the ovary (Bownes, 1986). In nematodes, synthesis takes place in the intestine of hermaphrodites (Kimble & Sharrock, 1983) while in echinoderms it occurs in the intestine and gonads of adults of both sexes (Shyu et al., 1986). Finally, in oviparous vertebrates, vitellogenin is produced under estrogen control in the liver of the adult female (Wallace, 1985; Wahli & Ryffel, 1985). In all species, newly synthesized vitellogenin is secreted into either the hemolymph, the coelomic fluid, or the bloodstream and is then selectively taken up by the growing oocytes in the ovary.

For several species, information has become available recently for both vitellogenin gene structure and expression. These include the invertebrates *Caenorhabditis elegans* (Spieth & Blumenthal, 1985; Spieth et al., 1985a) and *Drosophila melanogaster* (Barnett et al., 1980; Hung & Wensink, 1983; Yan et al., 1987) as well as the vertebrates *Xenopus laevis* (Wahli et al., 1979; Gerber-Huber et al., 1987) and the chicken *Gallus gallus* (Wang & Williams, 1983; van het Schip et al., 1987). In all these animals, vitellogenin is encoded in a small gene family, of which at least one member has been completely sequenced. The chicken and *Xenopus laevis* vitellogenins [200–250 kilodaltons (kDa)] are cleaved within the oocyte to give lipovitellin I (~120 kDa), the phosphoserine-rich phosvitin (~35 kDa), and lipovitellin II (~30 kDa). In *Caenorhabditis elegans*, there are two classes of yolk proteins. The first class consists of 170-kDa vitellogenins that are not cleaved and the second one of 115- and 88-kDa proteins which are derived from a short-lived 180-kDa vitellogenin precursor (Spieth & Blumenthal, 1985; Spieth et al., 1985a). In most insects, a 180–250-kDa precursor is cleaved into vitellins of 50–120 kDa (Bownes, 1986). Unfortunately, the amino acid sequence of such an insect vitellogenin is not yet available. One exception among insects to this general pattern is *Drosophila* in the sense that smaller uncleaved yolk proteins of 45 kDa are produced from three genes that have been sequenced (Barnett et al., 1980; Yan et al., 1987).

A previous comparison of the first 72 amino acids of the chicken major vitellogenin, the *Xenopus laevis* vitellogenin A2, and the *Caenorhabditis elegans* vitellogenin 5 (170-kDa class, see above) has suggested that the vitellogenic process in invertebrates and vertebrates is likely to be homologous rather than analogous (Spieth et al., 1985b). Since the complete primary structure of the *Xenopus* and chicken gene has recently been determined (van het Schip et al., 1987; Gerber-Huber et al., 1987), an extended comparison that defines the relatedness between these rather large proteins is now possible.

Experimental Procedures

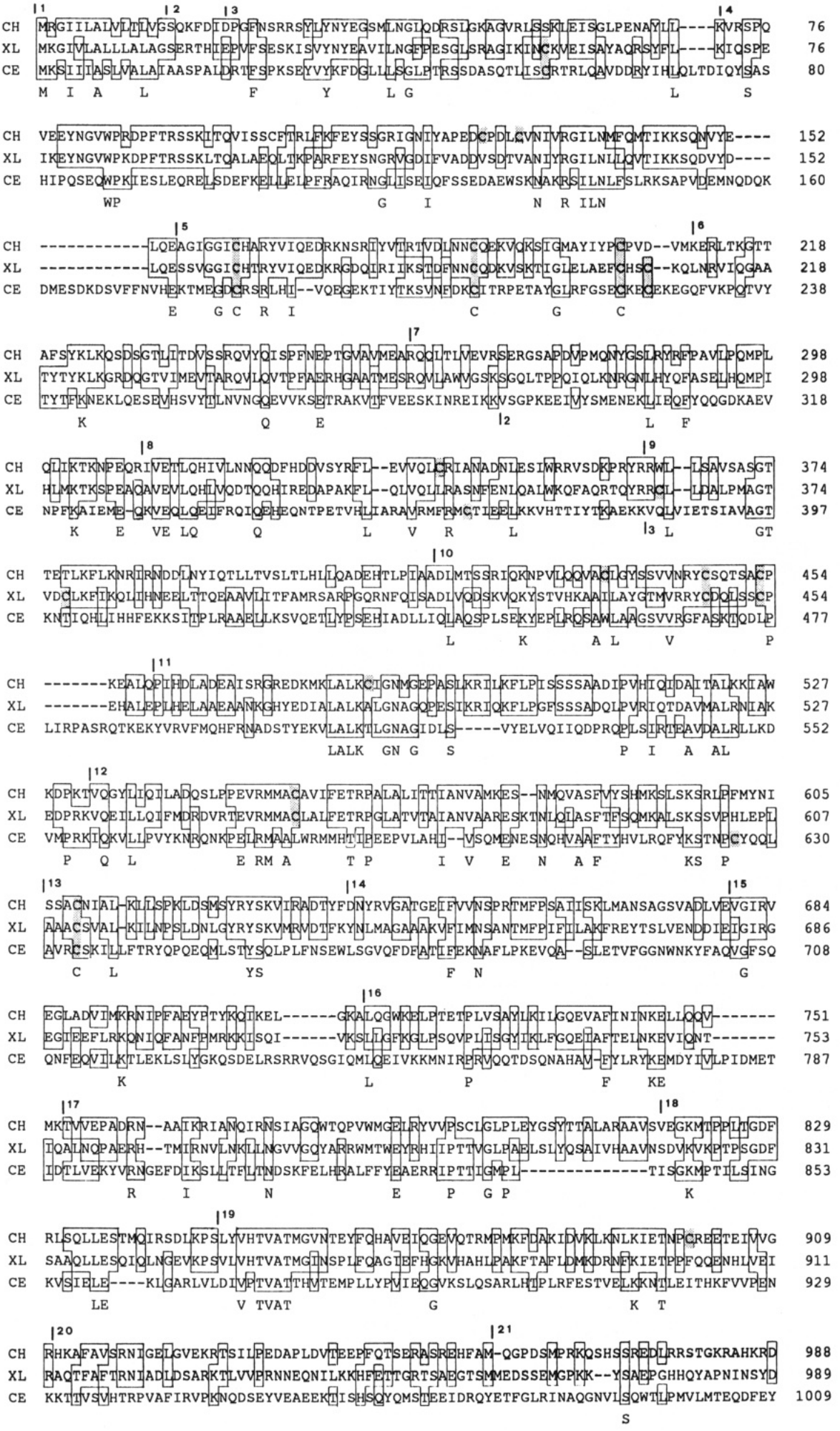
Computer Analysis of DNA and Protein Sequences. The analysis of the nucleotide sequences and the alignment of the pairwise-derived amino acid sequences of the *Caenorhabditis elegans* vitellogenin gene 5 (Spieth et al., 1985a), the *Xenopus laevis* vitellogenin gene A2 (Gerber-Huber et al., 1987), and the chicken major vitellogenin gene (van het Schip et al., 1987) have been performed with the program ANALYSIS (options 12 and 13), an adaptation for the Norsk computer (Rémy Fritz and Daniel Guinier, LGME, Strasbourg) of the program IDEAS (option SEQA, SEQF, SEOP; Minoru Kanehisa, NIH, Bethesda). The computer program used for the calculations and the graphic output of Figure 2 was written by Daniel Fahrni, IBA, Lausanne.

Results and Discussion

Homologies between Invertebrate and Vertebrate Vitellogenins. Figure 1 shows an alignment of the complete amino acid sequence of the chicken major vitellogenin, the *Xenopus laevis* vitellogenin A2, and the *Caenorhabditis elegans* vitellogenin 5 (170-kDa class; see the introduction), which unambiguously demonstrates the relatedness of these sequences. Furthermore, we observed some homology between the *Drosophila* and the three vitellogenins compared here, but we judged it to be too low to be shown. Clearly, a thorough evaluation of the relatedness between insect and other vitellogenins requires the elucidation of the primary structure of one of the common 180–250-kDa insect vitellogenins.

† This work was supported by the Swiss National Science Foundation, the Etat de Vaud, and the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).
‡ Université de Lausanne.
§ Present address: Institut für Biochemie und Molekularbiologie, CH-3012 Bern, Switzerland.
∥ University of Groningen.

0006-2960/87/0426-6397$01.50/0 © 1987 American Chemical Society


CH MRGIILALVLTLVGSQKFDIDPGFNSRRSYLYNYEGSMLNGIQDRSLGKAGVRLSSKLEISGLPENAYLL----KVRSPQ 76
XL MKGIVLALLLALAGSERTHIEPVFSESKISVYNYEAVILNGFPESGLSRAGIKINCKVEISAYAQRSYFL----KIQSPE 76
CE MKSIITASLVALAIAASPALDRTFSPKSEYVYKFDGLLLSGLPTRSSDASQTLISCRTRLQAVDDRYIHLQLTDIQYSAS 80
M I A L F Y L G L S
CH VEEYNGVWPRDPFTRSSKITQVISSCFIRLFKFEYSSGRIGNIYAPEDCPDLGVNIVRGILNMFQMTIKKSQNVYE---- 152
XL IKEYNGVWPKDPFTRSSKLTQALAEQLTKPARFEYSNGRVGDIFVADDVSDTVANIYRGILNILQVTIKKSQDVYD---- 152
CE HIPQSEQWPKIESLEQRELSDEFKELLELPFRAQIRNGLISEIQFSSEDAEWSKNAKRSILNLFSLRKSAPVDEMNQDQK 160
WP G I N R ILN
CH ----------LQEAGIGGICHARYVIQEDRKNSRIYVIRTVDLNNCQEKVQKSIGMAYIYPCPVD--VMKERLTKGTT 218
XL ----------LQESSVGGICHTRYVIQEDKRGDQIRIIKSTDFNNCQDKVSKTIGLELAEFCHSC--KQLNRVIQGAA 218
CE DMESDKDSVFFNVHEKTMEGDCRSRLHI--VQEGEKTIYTKSVNFDKCITRPETAYGLRFGSECKECEKEGQFVKPQTVY 238
E G C R I C G C
CH AFSYKLKQSDSGTLITDVSSRQVYQISPFNEPTGVAVMEARQQLTLVEVRSERGSAPDVPMQNYGSLRYRFPAVLPQMPL 298
XL TYTYKLKGRDQGTVIMEVTARQVLQVTPFAERHGAATMESRQVIAWVGSKSGQLTPPQIQLKNRGNLHYQFASELHQMPI 298
CE TYTFKNEKLQESEVHSVYTLNVNGQEVVKSETRAKVTFVEESKINREIKKVSGPKEEIVYSMENEKLIEQFYQQGDKAEV 318
K Q E L F
CH QLIKTKNPEQRIVETLQHIVLNNQQDFHDDVSYRFL--EVVQLCRIANADNLESIWRRVSDKPRYRRWL--LSAVSASGT 374
XL HLMKTKSPEAQAVEVLQHLVQDTQQHIREDAPAKFL--QLVQLLRASNFENLQALWKQFAQRTQYRRCL--LDALPMAGT 374
CE NPFKAIEME-GKVEQLQEIFRQIQEHEQNTPETVHLIARAVRMFRMCTIEELKKVHTTIYTKAEKKVQLVIETSIAVAGT 397
K E VE LQ Q L V R L L GT
CH TETLKFLKNRIRNDDLNYIQTLLTVSLTLHLLQADEHTLPIAADIMTSSRIQKNPVLQQVAQLGYSSVVNRYCSQTSACP 454
XL VDCLKFIKQLIHNEELTTQEAAVLITFAMRSARPGQRNFQISADLVQDSKVQKYSTVHKAAILAYGTMVRRYCDQLSSCP 454
CE KNTIQHLIHHFEKKSITPLRAAEPLLKSVQETLYPSEHIADLLIQLAQSPLSEKYEPIRQSAWLAAGSVVRGFASKTQDLP 477
L K A L V P
CH -------KEALQPIHDLADEAISRGREDKMKLALKCIGNMGEPASLKRILKFLPISSSSAADIPVHIQIDAITALKKIAW 527
XL -------EHALEPLHELAAEAANKGHYEDIALALKALGNAGQPESIKRIQKFLPGFSSSADQLPVRIQTDAVMALRNIAK 527
CE LIRPASRQTKEKYVRVFMQHFRNADSTYEKVLALKTLGNAGIDLS-----VYELVQIIQDPRQPLSIRTEAVDALRLLKD 552
LALK GN G S P I A AL
CH KDPKTVQGYLIQILADQSLPPEVRMMACAVIFETRPALALITTIANVAMKES--NMQVASFVYSHMKSLSKSRLPFMYNI 605
XL EDPRKVQEILLQIFMDRDVRTEVRMMACLALFETRPGLATVTAIANVAARESKTNLQIASFTFSQMKALSKSSVPHLEPL 607
CE VMPRKIQKVLLPVYKNRQNKPELRMAALWRMMHTIPEEPVLAHI--VSQMENESNQHVAAFTYHVLRQFYKSTNPCYQQL 630
P Q L E RM A T P I V E N A F KS P
CH SSACNIAL-KLLSPKLDSMSYRYSKVIRADTYFDNYRVGATGEIFVVNSPRTMFPSAIISKLMANSAGSVADLVEVGIRV 684
XL AAACSVAL-KILNPSLDNLGYRYSKVMRVDTFKYNLMAGAAAKVFIMNSANTMFPIFILAKFREYTSLVENDDIEIGIRG 686
CE AVRCSKILLFTRYQPQEQMLSTYSQLPLFNSEWLSGVQFDFATIFEKNAFLPKEVQA--SLETVFGGNWNKYFAQVGFSQ 708
C L YS F N G
CH EGLADVIMKRNIPFAEYPTYKQIKEL------GKALQGWKELPTETPLVSAYLKILGQEVAFININKELLQQV------- 751
XL EGIEEFLRKQNIQFANFPMRKKISQI------VKSLLGFKGLPSQVPLISGYIKLFGQEIAFTELNKEVIQNT------- 753
CE QNFEQVILKTLEKLSLYGKQSDELRSRRVQSGIQMLQEIVKKMNIRPRVQQTDSQNAHAM-FYLRYKEMDYIVLPIDMET 787
K L P F KE
CH MKTVVEPADRN--AAIKRIANQIRNSIAGQWTQPVWMGELRYVVPSCLGLPLEYGSYTTALARAAVSVEGKMTPPLTGDF 829
XL IQALNQPAERH--TMIRNVLNKLLNGVVGQYARRWMTWEYRHIIPTTVGLPAELSLYQSAIVHAAVNSDVKVKPTPSGDF 831
CE IDTLVEKYVRNGEFDIKSLLTFLTNDSKFELHRALFFYEAERRIPTTIGMPL--------------TISGKMPTILSING 853
R I N E P G P K
CH RLSQLLESTMQIRSDLKPSLYVHTVATMGVNTEYFQHAVEIQGEVQTRMPMKFDAKIDVKLKNLKIETNPCREETEIMVG 909
XL SAAQLLESQIQLNGEVKPSVLVHTVATMGINSPLFQAGIEFHGKVHAHLPAKFTAFLDMKDRNFKIETPPFQQENHLVEI 911
CE KVSIELE----KLGARLVLDIVPTVATTHVTEMPLLYPVIEQGVKSLQSARLHIPLRFESTVELKKNTLEITHKFVVPEN 929
LE V TVAT G K T
CH RHKAFAVSRNIGELGVEKRTSILPEDAPLDVTEEPFQTSERASREHFAM-QGPDSMPRKQSHSSREDLRRSTGKRAHKRD 988
XL RAQTFAFTRNIADLDSARKTLVVPRNNEQNILKKHFETTGRTSAEGTSMMEDSSEMGPKK--YSAEPGHHQYAPNINSYD 989
CE KKTTVSVHTRPVAFIRVPKNQDSEYVEAEEKTISHSQYQMSTEEIDRQYETFGLRINAQGNVLSQWTLPMVLMTEQDFEY 1009
S

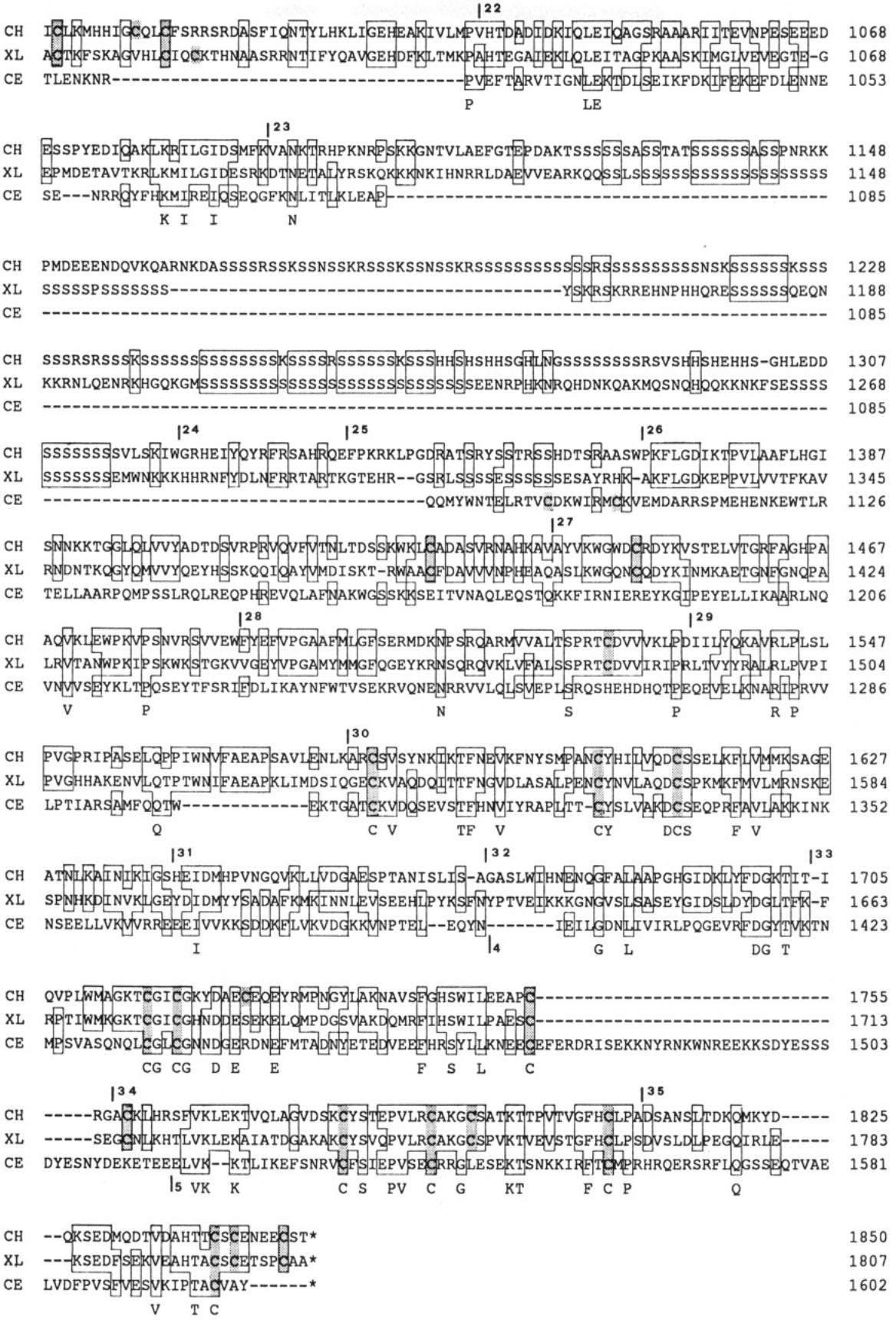


FIGURE 1: Comparison of the vitellogenin amino acid sequence (one-letter code) from chicken (CH; van het Schip et al., 1987), *Xenopus laevis* (XL; Gerber-Huber et al., 1987), and *Caenorhabditis elegans* (CE; Spieth et al., 1985a). Identical residues in two or in all three sequences are boxed, and those conserved in all three sequences are denoted underneath the sequences. The exon number and limits which are the same for the chicken and *X. laevis* gene are given over the CH sequence, and those of the *C. elegans* gene are given below the CE sequence. The cysteines are dashed. The cumulative number of residues of each of the three sequences is indicated at the right side of each line.

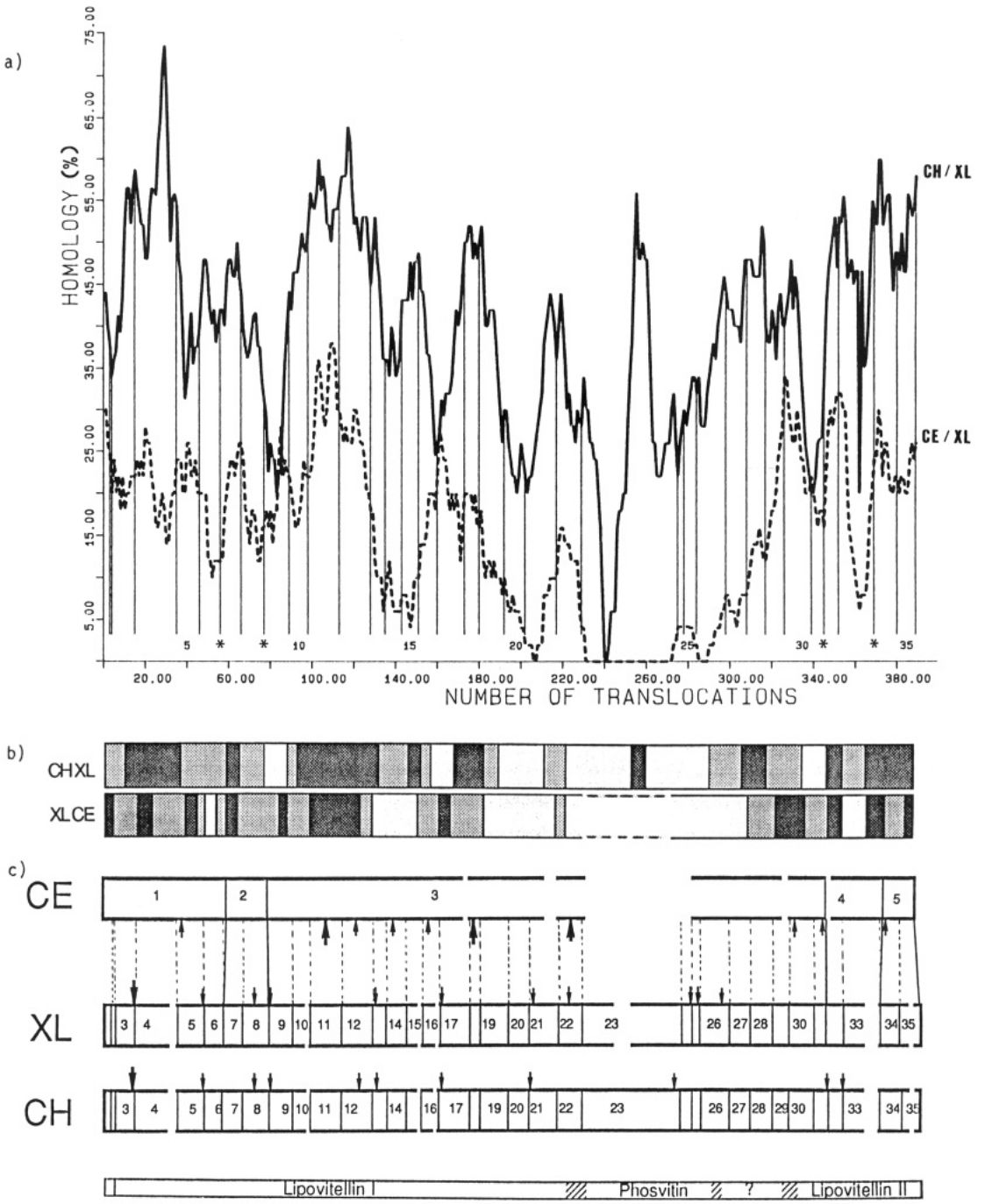


FIGURE 2: Variation of homology along the two pairs of amino acid sequences compared. (a) Plot of the percentage of identical residues between the chicken and *X. laevis* vitellogenins [CH/XL (—)] and between the *X. laevis* and *C. elegans* vitellogenins [XL/CE (- - -)]. The percent of homology (vertical axis) was calculated on the basis of the alignment shown in Figure 1, using a window of 50 residues translocated by 5 residues (1–50, 5–55, 10–60, etc.). The number of translocations is given on the horizontal axis. The exon junctions of the *X. laevis* and chicken genes are indicated by numbered vertical lines and those of the *C. elegans* gene by asterisks. The computer program used for these calculations has been written by Daniel Fahrni, IBA, Lausanne. (b) Schematic representation of the results shown in (a). In the CH/XL comparison, domains of homology are defined as follows: (heavy stipples) more than 45% homology; (light stipples) between 35% and 45% average homology; (no stipples) less than 35% homology. In the XL/CE comparison, heavy stipples represent more than 23% homology, light stipples are between 13% and 23% average homology, and no stipples are less than 13% homology. (c) Schematic representation of the exon organization of the *C. elegans* (CE), *X. laevis* (XL), and chicken (CH) vitellogenin genes as derived from Figure 1. Gaps of one and two amino acids (light downward arrow) and gaps of three to five amino acids (heavy downward arrow) are indicated at the locations they occur; longer gaps are given by an interruption, corresponding to their length, of the solid lines in the scheme. Exon junctions of the XL gene are projected onto the CE gene for comparison. The location of the vertebrate yolk proteins on the vitellogenin precursor is given. Striated areas indicate the limits of the different domains, and the question mark represents an area whose fate, during processing, is unknown as yet.

The *X. laevis* and chicken vitellogenins comprise 1808 and 1850 amino acids, respectively. The *C. elegans* vitellogenin is slightly shorter with 1603 amino acids. As shown in Figures 1 and 2, the degree of homology varies extensively along the amino acid sequences compared (chicken/*X. laevis* and *X. laevis*/*C. elegans*). The regions of higher homology most likely represent domains with specific functions which have yet to be identified. This is particularly illustrated by the conservation of the cysteine pattern, especially in the cysteine-rich lipovitellin II region (Figure 1).

A striking difference between the two vertebrate and the *C. elegans* vitellogenins is the absence of the phosvitin region

in the latter. Interestingly, this difference had been predicted from a comparison of the amino acid composition of the three proteins (Spieth et al., 1985a). The phosvitin domain is the least conserved between the two vertebrate vitellogenins, suggesting that it evolves rapidly (Figure 1). It might represent a region which specialized recently by duplications of serines or serine-rich stretches (Byrne et al., 1984). This is reflected by the clustering of this amino acid that represents up to 50% of the phosvitin domain and by the large size of the exon encoding this region (Wahli et al., 1979; Byrne et al., 1984). Alternatively, the phosphoserine domain might have been lost in invertebrates. The nature of the selective pressure favoring a phosphate and calcium carrier domain unique to vertebrates is unclear as yet. It might be linked to the formation of bones which starts long before hatching in vertebrates. However, this remains to be investigated, as well as the intervention, in general, of the different yolk constituents in developmental processes.

Evolution of the Intron–Exon Organization of the Vitellogenin Gene. The phylogenetic lines leading to nematodes and amphibians or birds have separated in the Precambrian. Selective pressure on the compared vitellogenins has maintained enough homology to reveal common ancestry. Although apparently quite different, the exon–intron pattern of the *C. elegans*, the *X. laevis*, and the chicken genes suggests common ancestry as well. The coding region of the *X. laevis* and chicken gene is interrupted 34 times by introns at corresponding positions. In contrast, the *C. elegans* gene has only 4 introns. However, exon boundaries of the *C. elegans* gene correspond exactly, or within a few amino acids, to exon boundaries of the vertebrate vitellogenin genes (Figures 1 and 2). The apparent conservation of these intron positions suggests that these introns were present in the ancestral gene. Limited intron sliding might have occurred where the boundaries do not coincide precisely. Consequently, the question arises whether the difference of 30 introns between the *X. laevis* or chicken gene and the *C. elegans* gene represents gains in the vertebrate gene or, alternatively, losses in the *C. elegans* gene, as compared to the ancestral vitellogenin gene. Part of the answer might come from the following observation. The alignment of the *C. elegans*, *X. laevis*, and chicken vitellogenin amino acid sequences was performed by computer. Interestingly, the gaps introduced by the computer in either of the three sequences often coincide, within a few residues, with exon junctions in the vertebrate genes or corresponding positions in the nematode gene. This might suggest rearrangements around these positions such as elimination of introns resulting in exon fusions in *C. elegans*. This interpretation is supported by the occurrence of unusually long exons in the *C. elegans* gene (Figure 2). Thus, the vertebrate gene may have more resemblance to the earliest vitellogenin gene than the nematode gene. A more definitive answer to the question will depend upon more information about the structure of other invertebrate and vertebrate vitellogenin genes. A comparative analysis extended to arthropodes, molluscs, echinoderms, and lower vertebrates (fishes, reptiles) will surely provide valuable data toward elucidating the mechanisms that have been involved in the evolution of the structural organization of the gene.

The discovery of introns in genes has generated different lines of thought (Doolittle, 1978; Darnell, 1978; Tiemeier et al., 1978; Crick, 1979; Gilbert, 1978). Particularly, their positions would reflect the evolutionary past of genes with respect to the assembly of structural and functional elements (Gilbert, 1979; Gilbert et al., 1986). However, prokaryotes and lower eukaryotes would have streamlined their genome by selection against the retention of introns which would represent a too heavy bioenergetic expense for rapidly multiplying cells (Darnell & Doolittle, 1986; Straus & Gilbert, 1985; Lonberg & Gilbert, 1985). Our results can be interpreted similarly. They suggest that some invertebrate genomes might have been streamlined compared to higher complexity genomes of vertebrates. This might be particularly true for the genome of the worm *Caenorhabditis* which, with 8×10^7 base pairs, is the smallest known among metazoan. Thus, the organization of the present day vitellogenin gene in vertebrates would be the most representative of the earliest gene.

ACKNOWLEDGMENTS

We thank Rêmy Fritz, Daniel Guinier (LGME, Strasbourg), and Daniel Fahrni (IBA, Lausanne) for computer programs, Igor Dawid, Bob Hipskind, Hans Trachsel, and Riccardo Wittek for critical reading of the manuscript, and Hannelore Pagel for typing the manuscript.

Registry No. Vitellogenin (chicken major reduced), 110026-29-4; vitellogenin (*Xenopus laevis* A2 reduced), 110026-30-7.

REFERENCES

Barnett, T., Pachl, C., Gergen, P. J., & Wensink, P. (1980) *Cell (Cambridge, Mass.) 21*, 729–738.

Bownes, M. (1986) *Annu. Rev. Entomol. 31*, 507–531.

Byrne, B. M., van het Schip, F. D., van de Klundert, J. A. M., Arnberg, A. C., Gruber, M., & AB, G. (1984) *Biochemistry 23*, 4275–4279.

Crick, F. (1979) *Science (Washington, D.C.) 204*, 264–271.

Darnell, J. E., Jr. (1978) *Science (Washington, D.C.) 202*, 1257–1261.

Darnell, J. E., & Doolittle, W. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1271–1275.

Doolittle, W. F. (1978) *Nature (London) 272*, 581–582.

Gerber-Huber, S., Nardelli, D., Haefliger, J.-A., Cooper, D. N., Givel, F., Germond, J.-E., Engel, J., Green, N. M., & Wahli, W. (1987) *Nucleic Acids Res. 15*, 4737–4760.

Gilbert, W. (1978) *Nature (London) 271*, 501.

Gilbert, W. (1979) in *Eukaryotic Gene Regulation: ICN–UCLA Symposia on Molecular and Cellular Biology* (Axel, R., Maniatis, T., & Fox, C. F., Eds.) pp 1–10, Academic Press, New York.

Gilbert, W., Marchionni, M., & McKnight, G. (1986) *Cell (Cambridge, Mass.) 46*, 151–154.

Hung, M.-C., & Wensink, P. C. (1983) *J. Mol. Biol. 164*, 481–492.

Kimble, J., & Sharrock, W. J. (1983) *Dev. Biol. 96*, 189–196.

Lonberg, N., & Gilbert, W. (1985) *Cell (Cambridge, Mass.) 40*, 81–90.

Shyu, A.-B., Raff, R. A., & Blumenthal, T. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3865–3969.

Spieth, J., & Blumenthal, T. (1985) *Mol. Cell Biol. 5*, 2495–2501.

Spieth, J., Denison, K., Zucker, E., & Blumenthal, T. (1985a) *Nucleic Acids Res. 13*, 7129–7138.

Spieth, J., Denison, K., Kirtland, S., Cane, J., & Blumenthal, T. (1985b) *Nucleic Acids Res. 13*, 5283–5295.

Straus, D., & Gilbert, W. (1985) *Mol. Cell Biol. 5*, 3497–3506.

Tiemeier, D. C., Tilghman, S. M., Polsky, F. I., Seidman, G. J., Leder, A., Edgell, M. H., & Leder, P. (1978) *Cell (Cambridge, Mass.) 14*, 237–245.

van het Schip, F. D., Samallo, J., Broos, J., Ophuis, J., Mojet, M., Gruber, M., & AB, G. (1987) *J. Mol. Biol.* (in press).

Wahli, W., & Ryffel, G. (1985) in *Oxford Surveys on Eukaryotic Genes* (Maclean, N., Ed.) Vol. 2, pp 96–120, Oxford University Press, Oxford, U.K.
Wahli, W., Dawid, I. B., Wyler, T., Jaggi, R. B., Weber, R., & Ryffel, G. U. (1979) *Cell (Cambridge, Mass.) 16*, 535–549.
Wallace, R. (1985) in *Developmental Biology* (Browder, L. W., Ed.) Vol. 1, pp 127–177, Plenum Press, New York.
Wang, S.-Y., & Williams, D. L. (1980) *Biochemistry 19*, 1557–1563.
Yan, Y. L., Kunert, C. J., & Postlethwait, J. H. (1987) *Nucleic Acids Res. 15*, 67–85.

Role of the Divalent Cation in Topoisomerase II Mediated Reactions†

Neil Osheroff
Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232
Received March 27, 1987; Revised Manuscript Received May 28, 1987

ABSTRACT: The effects of magnesium ions on interactions between *Drosophila melanogaster* topoisomerase II and its substrates were assessed by a number of kinetic and binding assays. Results indicated that the divalent cation plays two distinct functions in promoting enzyme–substrate interactions. One class of magnesium ions participates directly in enzyme-mediated DNA cleavage reactions. A second class of magnesium ions participates directly in topoisomerase II mediated ATPase reactions and functions by providing the enzyme with a magnesium–ATP substrate. In contrast, the divalent cation did not affect the quaternary structure of the enzyme, was not required for the site-specific binding of topoisomerase II to DNA, and did not affect the enzyme's ability to discern the topological state of its nucleic acid substrate.

Eukaryotic type II topoisomerases alter the topology of DNA by passing an intact helix of DNA through a transient, enzyme-bound, double-stranded break made in a second helix (Wang, 1982, 1985; Vosberg, 1985). This double-stranded DNA passage reaction proceeds catalytically at the expense of ATP hydrolysis and absolutely requires the presence of divalent magnesium ions (Wang, 1982, 1985; Vosberg, 1985).

Although many of the relationships between eukaryotic topoisomerase II and its substrates have been characterized, very little is known about the enzyme's requirement for magnesium. Many functions for the divalent cation are possible, including roles in enzyme subunit–subunit interactions, recognition and binding of nucleic acids by topoisomerase II, DNA cleavage reactions, DNA strand passage, ATP hydrolysis, and/or enzyme turnover. One previous study concluded that the enzyme's need for a divalent cation went beyond the possible requirement for a magnesium–ATP substrate (Osheroff et al., 1983). Other studies presented evidence which implicated magnesium in reaction steps which occurred during or before DNA cleavage (Sander & Hsieh, 1983; Liu et al., 1983; Pommier et al., 1984). Unfortunately, while the above work confirmed that magnesium was required for interactions between topoisomerase II and its substrates (i.e., DNA and ATP), no specific role for the divalent cation could be ascribed.

In order to more fully define the requirement of *Drosophila melanogaster* topoisomerase II for magnesium, the effects of this divalent cation on the interactions between the enzyme and its substrates were assessed by a number of kinetic and binding assays. Results of the present work indicate that magnesium plays two distinct functions in topoisomerase II–substrate interactions. One class of magnesium ions participates directly in enzyme-mediated DNA cleavage reactions. A second class of magnesium ions promotes topoisomerase II mediated ATPase activity by fulfilling the requirement for a magnesium–ATP substrate.

A preliminary account of some of this work has appeared (Osheroff, 1985).

Experimental Procedures

D. melanogaster DNA topoisomerase II was purified from the nuclei of Kc tissue culture cells or 6–18-h-old embryos by the procedure of Shelton et al. (1983). Negatively supercoiled bacterial plasmid pBR322 (Bolivar et al., 1977) DNA was isolated from *Escherichia coli* DH1 (Hanahan, 1983) by a Triton X-100 lysis procedure followed by double banding in cesium chloride–ethidium bromide gradients (Maniatis et al., 1982). Analytical reagent-grade $MgCl_2 \cdot 6H_2O$ was obtained from Fisher; ethidium bromide and tris(hydroxymethyl)-aminomethane (Tris)[1] were from Sigma; SDS and proteinase K were from E. Merck Biochemicals; bovine serum albumin (nuclease free) was from BRL; and adenosine 5′-[γ-^{32}P]triphosphate (3000 Ci/mmol) and ACS aqueous counting scintillant were from Amersham. All other chemicals were analytical reagent grade.

Agarose Gel Electrophoresis and Quantitation of Reaction Products. With the exception of ATPase experiments, the results of assays described below were assessed by agarose gel electrophoresis. After reactions were completed, products were mixed with loading buffer (60% sucrose, 0.05% bromophenol blue, 0.05% xylene cyanole FF, and 10 mM Tris-HCl, pH 7.9) and applied to 1.0% agarose (MCB) gels. Samples (0.3 μg of DNA) were subjected to electrophoresis in 40 mM Tris–acetate, pH 8.3, and 2 mM EDTA at 5 V/cm. Gels were stained for 30 min in an aqueous solution of ethidium bromide (1 μg/mL). DNA bands were visualized by transillumination

† This work was supported by National Institutes of Health Research Grant GM-33944 and Biomedical Research Support Grant RR-05424.

[1] Abbreviations: Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; SDS, sodium dodecyl sulfate; EDTA, ethylenediaminetetraacetic acid.

0006-2960/87/0426-6402$01.50/0 © 1987 American Chemical Society