Rudimentary Phosvitin Domain in a Minor Chicken Vitellogenin Gene†

B. Marion Byrne,‡ Harry de Jong,‡ Ronaldus A. M. Fouchier,‡ David L. Williams,§ Max Gruber,‡ and Geert AB*,‡

Biochemisch Laboratorium, Groningen University, Nijenborgh 16, 9747 AG Groningen, The Netherlands, and Department of Pharmacological Sciences, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794-8651

Received August 18, 1988; Revised Manuscript Received October 12, 1988

ABSTRACT: We have determined the nucleotide sequence and the derived amino acid sequence of the phosphoprotein-encoding region of the chicken vitellogenin III gene. The sequence of this minor vitellogenin could be aligned with exon 22 up to exon 27 of the previously sequenced major vitellogenin II gene (van het Schip et al., 1987). The exon 23 and 25 sequences are rich in serine codons (26% and 41%, respectively), and this region encodes at least one of the small egg yolk phosphoproteins. The major egg yolk phosphoprotein, phosvitin, is encoded by the analogous region in vitellogenin II. Comparison of the vitellogenin II and vitellogenin III sequences shows a great reduction in the size of the putative exon 23 of the latter (321 base pairs as opposed to 690). The number of serine codons is also drastically reduced from 124 in exon 23 of the vitellogenin II gene to 28 in vitellogenin III. The grouping of synonymous serine codons, as has hitherto been observed in sequenced vitellogenin phosphoproteins, has been maintained in vitellogenin III. A putative asparagine-linked N-glycosylation site which was conserved in the chicken vitellogenin II and the *Xenopus laevis* vitellogenin A2 gene, at the beginning of exon 23, is also present in vitellogenin III. The two chicken vitellogenins show a low conservation in the phosphoprotein-encoding region (average 33%, at the protein level) compared to that in the peripheral sequences (58% identity), which indicates that it is a rapidly evolving domain of the vertebrate vitellogenin gene.

Vitellogenin (Vtg)[1] is widespread throughout the oviparous animal kingdom making it an interesting candidate for the study of the molecular processes of evolution [for a recent review, see Wallace (1985)]. Tissue-specific and estradiol-induced synthesis of this yolk protein precursor occurs in the liver of egg-laying vertebrates (Gruber et al., 1976). The vitellogenin polypeptides are transported in the blood to the oocyte where they are cleaved into the major yolk proteins lipovitellin I and II (LvI and LvII) and the phosphoproteins phosvitin and phosvette (Bergink et al. 1974; Wiley & Wallace, 1981).

The *Xenopus laevis* and chicken vitellogenins have been studied in detail at the molecular level. In *X. laevis* four vitellogenin genes have been identified: A1, A2, and B1 linked on one chromosome and the closely related B2 gene at a separate genetic locus (Wahli et al., 1982; Schubiger & Wahli, 1986). The A2 gene has been sequenced entirely and encodes LvI, LvII, and either one phosvitin or two phosvettes (Gerber-Huber et al., 1987). In chicken, three vitellogenins, namely, VtgI, VtgII, and VtgIII, have been characterized (Wang & Williams, 1980; Wang et al., 1983). They accumulate in the plasma of roosters, that have received exogenous doses of estradiol, to a ratio of 0.33:1.0:0.08, respectively. Amino acid analyses indicated the presence of a highly phosphorylated serine-rich phosvitin in VtgI and II (116 mol of P each) and of a low molecular weight phosphopeptide (phosvette) in VtgIII (44 mol of P). There are at least five phosphoproteins in chicken egg yolk; two phosvitins, the major and the minor one (Clark, 1970; Wicks & Clark, 1987); and three phosvettes, E1, E2, and F (Wallace & Morgan, 1986). Only the VtgII gene has been fully characterized and sequenced (van het Schip et al., 1987). It encodes LvI and LvII and the major chicken phosvitin (Byrne et al., 1984). Recently, Evans et al. (1988) isolated cDNA clones corresponding to VtgI and VtgIII.

When the *X. laevis* A2 gene and the chicken VtgII gene were compared, the length and number of exons were highly conserved whereas the intron sequences showed no similarity whatsoever (Nardelli et al., 1987a). The 35 exons have an average of 40% identical amino acids and 52% identical nucleotides. In the phosphoprotein-encoding regions only 23% identity at the amino level and 34% at the nucleotide level was found, meaning that this region has evolved at a much faster rate than the rest.

Phosvitin is the most notable yolk polypeptide derived from the Vtg precursor because of its extraordinary amino acid composition. More than half of its residues are serines, most of which are phosphorylated. The high phosphorylation status (10% P) is believed to be of importance in sequestering calcium, iron, and other cations for the developing embryo (Taborsky, 1980). The size of the phosvitin moiety increases from lower to higher vertebrates as a consequence of an increased serine content. It has been proposed that the increase in phosvitin size may be required for embryonic bone formation in the higher nonmammalian vertebrates (Nardelli et al., 1987a). In agreement with this is the finding that precisely the phosvitin domain is missing from nematode Vtg genes (Spieth et al., 1985). Previous evidence suggested that the phosvitin moiety is involved in the receptor-mediated uptake

† This work was supported by the Netherlands Foundation for Chemical Research (Scheikundig Onderzoek Nederland) with financial aid from the Netherlands Organization for Scientific Research (Nederlands Organisatie voor Wetenschappelijk Onderzoek) and in part by National Institutes of Health Grant DK18171.

* Address correspondence to this author.

‡ Groningen University.

§ State University of New York at Stony Brook.

[1] Abbreviations: kb, kilobase (pair); bp, base pair(s); Vtg, vitellogenin; Lv, lipovitellin; cDNA, DNA complementary to mRNA; Py, pyrimidine; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetate.

0006-2960/89/0428-2572$01.50/0

of Vtg by oocytes (Yusko et al., 1981; Miller et al., 1981); however, Stifani et al. (1988), who recently solubilized and characterized the chicken oocyte Vtg receptor, disagree with the previous findings. Knowledge of the structure of phosvitin at the molecular level may help elucidate its exact role in embryogenesis.

We are interested in the relationship between members of the chicken Vtg family and in partiular in the evolution of the phosvitin domain (Byrne et al., 1984). We have used anti-VtgIII sera to screen a chicken liver cDNA library in the expression vector λ gt11 (Young & Davis, 1983). Here we present the nucleotide sequence and derived amino acid sequence of a cDNA clone which reacted with anti-VtgIII serum and of the corresponding genomic frament which encodes the phosphoprotein region of VtgIII. In addition, we compare this sequence to the phosvitin sequence of VtgII, which was determined previously (Byrne et al., 1984).

Experimental Procedures

Construction of a Chicken Liver cDNA Library. Polysomal RNA was prepared from estrogen-stimulated rooster liver (AB et al., 1976). The poly(A+) fraction was obtained by oligo(dT)–cellulose chromatography (Aviv & Leder, 1972). Single-stranded cDNA was synthesized by reverse transcriptase (Life Sciences Inc., St Petersbourg, FL) with oligo$(dT)_{12-15}$ as a primer. Second strand synthesis was performed with RNase H (Bethesda Research Laboratories, Gaithersburg, MD) and *Escherichia coli* DNA polymerase I, endonuclease-free (Boehringer Mannheim), as described by Gubler and Hoffman (1983). The double-stranded cDNA was prepared for cloning in λ gt11 essentially according to Huynh et al. (1984). After the ligation to phosphorylated *Eco*RI linkers, the excess linkers were removed by digestion with *Eco*RI and 2-propanol precipitation. The cDNA was ligated to *Eco*RI-cut dephosphorylated arms of λ gt11 phage (Protoclone, Promega Biotec, Madison, WI). Plaques were screened with anti-VtgIII serum prepared as previously (Wang et al., 1983) and diluted 1:1600 for the immunodetection of positive clones largely as described by Huynh et al. (1984). Goat anti-rabbit alkaline phosphatase conjugate was the second antibody, and nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) were color substrates (Protoblot, Promega Biotec, Madison, WI). After purification of positive clones, phage DNA was isolated from liquid cultures according to Maniatis et al. (1982). Phage DNA was digested with *Eco*RI, and the inserts were purified from an 0.8% low melting point agarose gel.

Northern and Southern Blotting. Poly(A+) RNA from an estradiol-treated rooster and from a control bird was prepared as for cDNA synthesis. The RNA was glyoxylated and separated on a 1% agarose gel (McMaster & Carmichael, 1977) and transferred to a Nytran 0.45-μm nylon membrane (Schleicher & Schuell, Dassel, FRG) which was presoaked for 5 min in distilled water. After blotting, the membrane was rinsed briefly in 50 mM Tris-HCl (pH 8.3), 50 mM boric acid, and 0.1 mM EDTA, dried, and UV irradiated for 8 min.

A Southern blot of chicken genomic DNA digested with *Eco*RI was performed with nylon filters as described above.

Cloning of a Genomic 3.6-kb EcoRI Fragment. Chicken DNA was digested with *Eco*RI and resolved on an 0.8% agarose gel. The 3.6-kb region was excised, and DNA was purified by isotachophoresis (Öfverstedt et al., 1984). The DNA was cloned in the *Eco*RI site of dephosphorylated λ gt10 arms (Promega Biotec, Madison, WI) and plated out on *E. coli* C600Hfl (Huynh et al., 1984). By use of a *Pst*I clone from the exon 21–22 region of VtgII as a probe (van het Schip et al., 1987), 2×10^5 clones were screened. Sixteen positive clones were isolated, and the 3.6-kb insert was purified out of one.

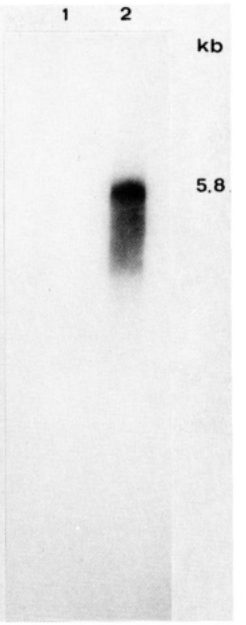

FIGURE 1: Northern blot analysis of a VtgIII cDNA clone. RNAs were extracted from the livers of a normal rooster (lane 1) and a rooster 72 h afer estradiol treatment (lane 2). About 10 μg of poly(A+) RNA was loaded in each lane. The M13 clone containing VtgIII cDNA as insert was used as probe.

DNA Sequencing. The cDNA phage inserts were subcloned in the *Eco*RI site of M13mp9 (Messing & Vieira, 1982). The 3.6-kb genomic fragment was first digested with *Pst*I and *Sau*3AI and subcloned in suitable M13 vectors. Sequencing was performed by the dideoxynucleotide chain termination method (Sanger et al., 1977). The alignment of the nucleotide sequences was calculated with the program IALIGN (Dayhoff et al., 1983). Minor adjustments in the alignment were sometimes necessary in order to align the derived amino acid sequence and the analogous exons.

Results and Discussion

Isolation and Identification of VtgIII cDNA Clones. A λ gt11 chicken liver cDNA library containing 10^5 clones was screened with anti-VtgIII serum (Wang et al., 1983) and yielded 17 positive clones. The insert of one of these encoded a short serine-rich region having some similarities to the VtgII phosvitin sequence and was 413 bp long. The cDNA clone hybridized to a single mRNA species of approximately 5.8 kb (Figure 1), which is the expected size for a Vtg mRNA. A signal was picked up exclusively in the poly(A+) fraction from an estradiol-stimulated rooster, and not from a control rooster. With the cDNA clone as probe in a Southern blot with chicken genomic DNA digested with *Eco*RI, a 3.6-kb band hybridized (Figure 2a).

A 3.6-kb genomic fragment of the VtgIII gene was isolated following an independent approach in which a search was made for related Vtg genes on the basis of cross-hybridization with VtgII probes. When a Southern blot of *Eco*RI-digested chicken DNA was probed with a *Pst*I fragment from the exon 21/22 region of VtgII, a very faint 3.6-kb band hybridized in addition to the expected band of 5.9 kb (Figure 2b). We excised the 3.6-kb fraction and cloned it in λ gt10 (Huynh et al., 1984). We screened 2×10^5 clones and found 16 positives, using the same *Pst*I fragment as mentioned above as probe. Preliminary evidence that the 3.6-kb *Eco*RI fragment originated from the VtgIII gene was obtained from Southern blot

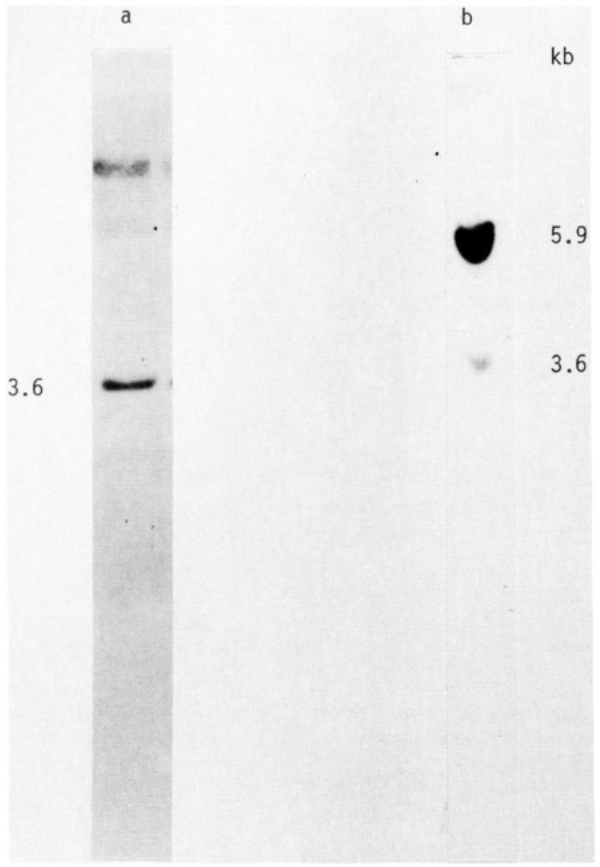

FIGURE 2: Southern blot analyses performed with 10 μg of chicken genomic DNA which was digested with *Eco*RI, electrophoresed on two separate 0.8% agarose gels, and blotted. (a) The blot was hybridized to the M13 clone containing VtgIII cDNA as insert. (b) The probe used was a *Pst*I fragment from the exon 21/22 region of the VtgII gene (van het Schip et al., 1987).

analysis, which yielded a strong signal with the VtgIII cDNA clone (data not shown). Sequence analysis of one of the genomic clones and comparison to the VtgIII cDNA clone proved that we were dealing with one and the same gene.

Nucleotide Sequence and Derived Amino Acid Sequence. The sequencing strategy is outlined in Figure 3. The partial nucleotide sequence of the 3.6-kb genomic fragment and the derived amino acid sequence are presented in Figure 4. VtgIII sequences resembling the region between exons 22 and 27 of the VtgII gene were identified. The exact positions of the putative exon 22/23, 23/24, and 24/25 junctions were mapped by comparison with the cDNA sequence. The position of the remaining splice junctions was predicted by comparison with the VtgII sequence, taking the consensus sequence for splice junctions into account (Mount, 1982). Table I gives a comparison of the exon lengths and of the percentage identity of the sequenced part of VtgIII with VtgII. It should be emphasized that the exon numbering for the VtgIII gene is purely tentative and is based upon the strong conservation in number and position of introns interrupting the coding region (Nardelli et al., 1987a). Whereas exons 26 and 27 have identical lengths and exons 24 and 25 differ by only one triplet, exon 23 is 369 nucleotides shorter in VtgIII than in VtgII. Exons 23, 24, and 25 show only a 36%, 38%, and 24% identity, respectively, at the amino acid level and 55%, 52%, and 46% at the nucleotide level. Exons 26 and 27 show a 59% and 57% identity at the amino acid level, respectively, and 66% at the nucleotide level between the aligned chicken sequences. On the whole, the conservation is rather low for homologous genes within a species; however, the moderate similarity at the nucleotide level confirms our findings in Southern blot hybridization studies. VtgII probes picked up only a very weak band which could be assigned to the VtgIII gene (Figure 2b).

Table I: Comparison of Exons 23–27 of Both Chicken Vitellogenins[a]

	exon length (bp)		no. of serine residues		% identity	
exon	VtgIII	VtgII	VtgIII	VtgII	NT	AA
23	321	690	28	124	55	36
24	48	51	0	1	52	38
25	87	90	12	7	46	24
26	213	213	8	5	66	59
27	143	143	3	3	66	57

[a] Comparison of the exon lengths, the number of serine residues per exon, and the percentage identity at nucleotide (NT) and amino acid (AA) levels are on the basis of the alignment of Figure 4.

Phosphoprotein-Encoding Region. The most conspicuous difference between the two sequences lies in their serine content (Table I). Exon 23 of the VtgII gene encodes 124 serines, and the analogous exon in VtgIII encodes only 28 whereas exon 25 of the VtgIII gene encodes 12 serine residues in contrast to 7 in the VtgII gene. The phosvitin of VtgII is encoded mainly by exon 23, and seven amino acids at the C-terminus are derived from exon 24. The N-terminal sequence of the minor, methionine-free phosvitin of hen egg yolk has been analyzed (Clark, 1985) and is presumably encoded by VtgI since it does not coincide with the present sequence nor with the VtgII sequence. We conclude that our sequence is derived from the VtgIII gene since we picked up the cDNA clone with a polyclonal anti-VtgIII serum which was not cross-reactive with other Vtgs (Wang et al., 1983). Furthermore, the number of serines calculated from the derived amino acid sequence for the phosphoprotein-encoding region coincides with previous predictions. Wang et al. (1983) already predicted that VtgIII, with 44 mol of P, would contain a phosvette, a smaller phosphoglycoprotein, rather than a

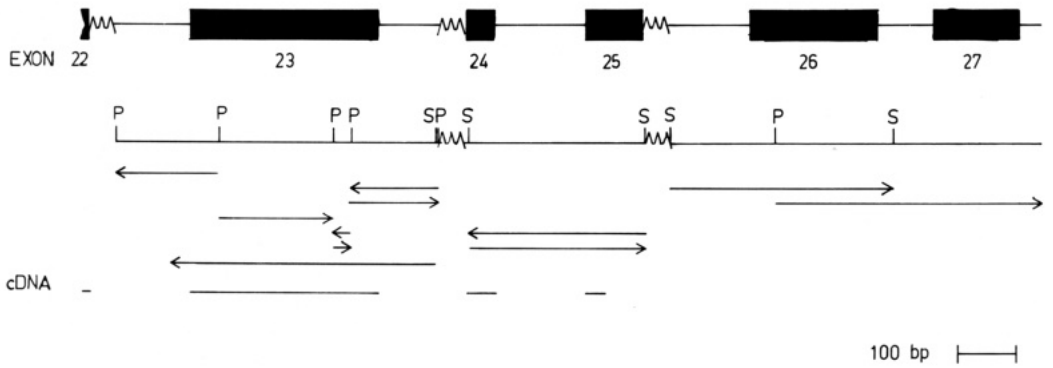

FIGURE 3: Sequencing strategy. Exons are shown at the top (black boxes) with their putative numbers underneath. Jagged lines denote gaps in the (intron) sequences. The cDNA was sequenced in both directions. P (*Pst*I); S (*Sau*3AI).

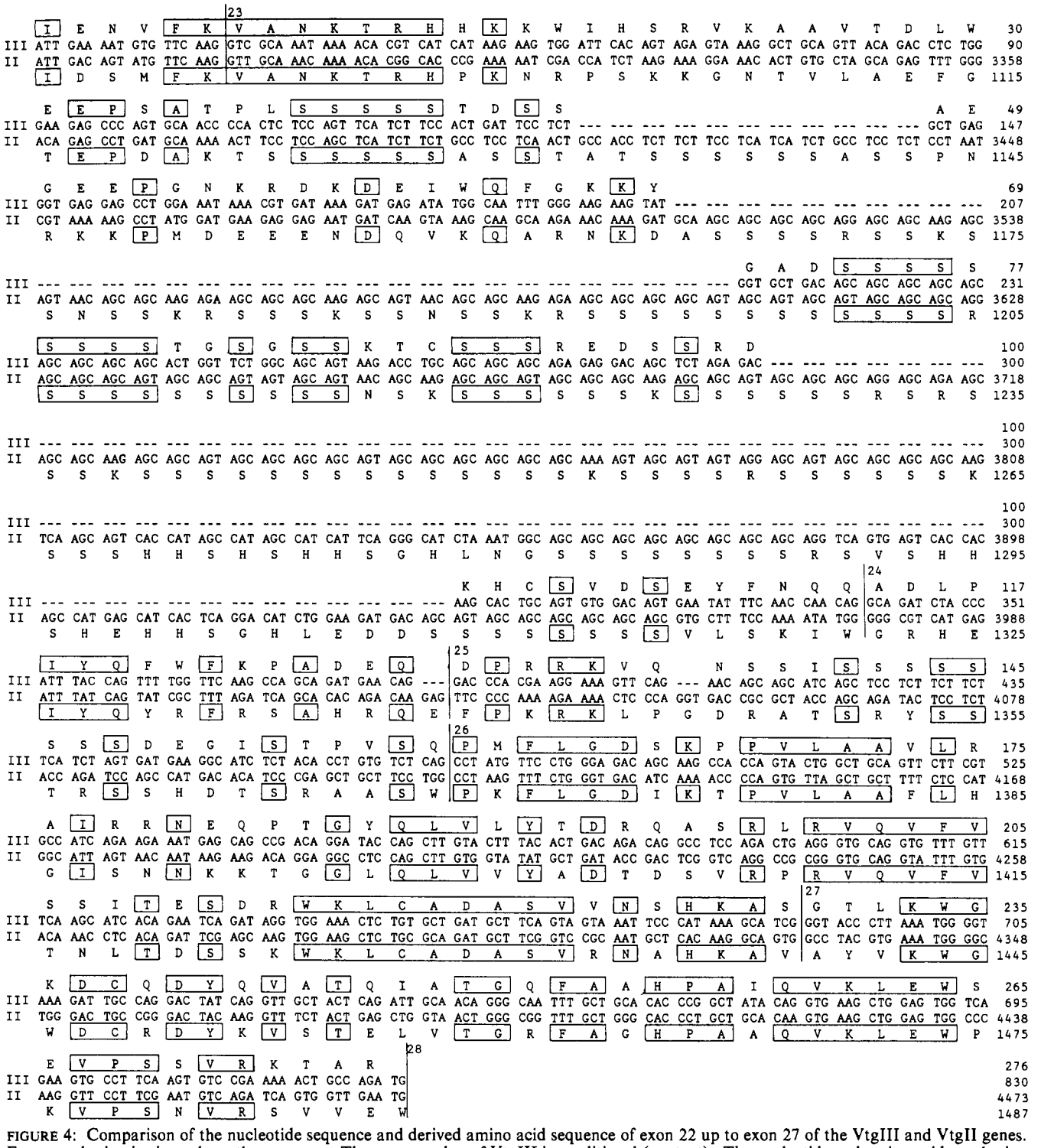

FIGURE 4: Comparison of the nucleotide sequence and derived amino acid sequence of exon 22 up to exon 27 of the VtgIII and VtgII genes. Exon numbering is given above the sequence. The exon number of VtgIII is conditional (see text). The nucleotide and amino acid numbering of both sequences is given at the right. The numbering for VtgIII is purely arbitrary starting with 1. The numbering for VtgII is taken from Nardelli et al. (1987). Identical amino acids are boxed.

larger phosvitin (Clark, 1970). Since most or all of the phosphates are attached to serine residues in the phosphopeptide domain, one would expect a total of approximately 44 serines in the VtgIII phosvette. Exons 23/25 taken together contain 40 serines, which coincides rather well with the predicted amount. These data provide direct evidence that the VtgIII polypeptide is encoded by a separate Vtg gene and is not, for instance, the result of alternative splicing of the VtgII gene.

Recently, Wallace and Morgan (1986) separated three small phosphoproteins, which they termed phosvettes, from hen egg yolk and concluded that either phosvette E1 (M_r 15 000) or phosvette E2 (M_r 18 000) is derived from VtgIII. Exact molecular weight determination of these highly phosphorylated proteins has proven difficult by conventional procedures, and these data may be overestimated (Wallace & Morgan, 1986). Neither the N- or C-terminal sequence nor the amino acid composition of the phosvettes is known. However, the sequence IYQ which terminates the VtgII phosvitin sequence (Byrne et al., 1984) had been conserved (position 1328) and may well be used in VtgIII also. The N-terminus may lie at the same position as in VtgII phosvitin making TDLWEEP the N-

terminal peptide which has close similarity to the N-terminal phosvitin sequence TELGTDP of the duck (Clark, 1985). If we choose these hypothetical N- and C-termini, the protein formed would have a M_r of 11 807 which, with added phosphate (27 P) would amount to M_r 13 967. Exon 25 of VtgIII, with its high number of serine codons, may encode a separate phosvette. This is the first time that a relationship has been established between the small egg yolk phosphoproteins or phosvettes and a Vtg gene.

The VtgII phosvitin domain was confined mainly to exon 23 whereas *X. laevis* has an extra serine-rich region in exon 25. In this respect, the situation in VtgIII resembles that of the *X. laevis* VtgA2 gene where there has been a preferential amplification of serine codons in exon 23 as well as in exon 25.

Taking the threonine and alanine residues at positions 27 and 1112 of the VtgIII and VtgII sequences, respectively, as the N-terminus of the phosphoprotein, the sequences can be aligned up to the amino acid at position 69 of VtgIII. There are few conserved residues in this region. The two conserved proline residues (positions 33 and 53) form β-turns in the model for the secondary structure of the VtgII phosvitin proposed by Renugopalakrishnan et al. (1985). β-Turns are often situated near serine phosphorylation sites (Small et al., 1977). The conserved aspartate (position 60), glutamine (position 64), and lysine (position 68) residues are situated on one side of an α-helix in the proposed secondary structure model and may therefore be of functional significance. The entire "core" region of mainly serines in exon 23 of VtgII (positions 1168–1265) is absent as such from the VtgIII sequence. It is replaced by a block of 31 residues in VtgIII (positions 70–100), 17 of which are serines. The remaining sequence up to the C-terminus of the VtgII phosvitin (position 1328) displays little similarity either at the amino acid level or at the nucleotide level. There are two cysteine residues in this region of the VtgIII sequence (positions 90 and 103).

Serine Codon Usage. The codon usage for the groups of serines within the three vertebrate Vtgs (chicken VtgII and III and *X. laevis* A2) sequenced thus far is remarkably alike. The "head group" of serines at the 5′ end of exon 23 are encoded chiefly by TCX codons and this sequence is duplicated in the VtgII and A2 genes only. The serines in the core region are encoded by AGPy. The serines in exon 25 are encoded by both types of codons although mainly TCX. We presume that numerous unequal crossing-over events occur between the stretches of synonymous serine codons thereby allowing expansions of these regions.

Cysteine Motifs. Exon 26 and 27 contain a number of homologous blocks of amino acid residues. In particular, the two cysteines in these exons are conserved in the VtgII and VtgIII genes. A block of nine residues around the cysteine at position 217 is 100% conserved. All four cysteines in the VtgIII sequence occur in the triplet KXC with an acidic residue in close proximity. The lipovitellin II region of the VtgII and the A2 genes and even the *Caenorhabditis elegans* Vtg gene is rich in (conserved) cysteine residues (Nardelli et al., 1987b). These sequences may be essential in the formation of disulfide bridges for proper folding of the protein. The conformational structure of Vtg is likely to be important for its transport as a dimer by the blood, its uptake by the oocyte and subsequent proteolytic cleavage, and the storage of its protein products, Lv and phosvitin, in the (semi)crystalline state as granules or platelets in the egg yolk (Wallace, 1985).

Glycosylation. The VtgII phosvitin is glycosylated at position 1280 (Shainkin & Perlmann, 1971; Byrne et al. 1984), but this site is lacking in the VtgIII sequence. However, the sequence FKVANKTRH at the exon 22/23 junction, which includes a potential asparagine-linked N-glycosylation site (Marshall, 1972), is homologous in both chicken Vtgs (Figure 4). This potential glycosylation site at the beginning of exon 23 is also the only site conserved between the chicken VtgII and the *X. laevis* Vtg A2 gene. Consequently, it may indeed by glycosylated in vivo and have a special function in the Vtgs. Carbohydrate has been implicated in the ovarian uptake of yolk proteins (Miller et al., 1981) and in the cellular recognition of other glycoproteins [for review, see Neufield and Ashwell (1980)]. We postulate that the putative N-glycosylation site right beside the phosphoprotein-coding region at the beginning of exon 23 may function in the uptake or endocytosis of Vtg by the oocyte.

Evolution. Taking an average of 5% divergence per 30 million years (outside of the phosphoprotein-encoding exons), we can conclude from the present data that the chicken VtgII and VtgIII genes diverged about 240 million years ago. The *X. laevis* A–B gene pair (80% identity) originated about 150 million years ago as a result of a gene duplication (Germond et al., 1984) whereas the A1 and A2 type (95% identity) and B1 and B2 type genes (95% identity) evolved more recently, namely, some 30–40 million years ago (Schubiger & Wahli, 1986). Therefore, the chicken and *Xenopus* Vtg gene families are the result of independent gene and genome duplication events since the amphibian/reptile/bird divergence some 350 million yeas ago. A detailed comparison of more diverse Vtg gene sequences is needed before the exact evolutionary origins of each gene can be elucidated. Clearly the phosphoprotein-encoding domain of the Vtg genes is a hypervariable region. It has grown and expanded with time from lower to higher vertebrates at a much faster rate than the rest of the gene, as a result of numerous unequal crossing-over events, which would be advantageous in its presumed role as a phosphate and calcium source in embryonic bone formation. By comparison studies we have found very few homologous sequences in the different Vtg phosphoproteins sequenced to date. Clearly, the serines are the essential component. It will be interesting to elucidate and compare the phosvitin/phosvette sequences of more diverse species such as reptiles and fish in the future.

ACKNOWLEDGMENTS

We thank Dr. Peter Terpstra for assistance in the computer analysis and John Samallo for providing M13 vectors. We acknowledge the assistance of Marco Harmsen, AnneMarie van Leeuwen, and Monique den Heijer at various stages of this work. We are grateful to Dr. Ann Gibbins and Dr. Jaap Beintema for the stimulating discussions.

REFERENCES

AB, G., Roskam, W. G., Dijkstra, J., Mulder, J., Willems, M., van der Ende, A., & Gruber, M. (1976) *Biochim. Biophys. Acta 454*, 67–78.

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

Bergink, E. W., Wallace, R. A., Van den Berg, J. A., Bos, E. S., Gruber, M., & AB, G. (1974) *Am. Zool. 14*, 1177–1193.

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

Clark, R. C. (1970) *Biochem. J. 118*, 537–542.

Clark, R. C. (1985) *Int. J. Biochem. 17*, 983–988.

Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983) *Methods Enzymol. 91*, 524–545.

Evans, M. I., Silva, R., & Burch, J. B. E. (1988) *Genes Dev. 2*, 116–124.
Gerber-Huber, S., Nardelli, D., Haefliger, J.-A., Cooper, D. N., Givel, F., Germond, J.-E., Engel, J., Green, M., & Wahli, W. (1987) *Nucleic Acids Res. 15*, 4737–4760.
Germond, J.-E., Walker, P., ten Heggeler, B., Brown-Luedi, M., de Bony, E., & Wahli, W. (1984) *Nucleic Acids Res. 12*, 8595–8609.
Gruber, M., Bos, E. S., & AB, G. (1976) *Mol. Cell. Endocrinol. 5*, 41–50.
Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.
Huynh, T. V., Young, R. A., & Davis, R. W. (1984) in *DNA Cloning Techniques: A Practical Approach* (Glover, D., Ed.) pp 98–121, IRL Press, Oxford.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning, a Laboratory Manual*, pp 76–85, Cold Spring Harbor Laboratory, Cold Spring Habor, NY.
Marshall, R. D. (1972) *Annu. Rev. Biochem. 41*, 673–702.
McMaster, G. K., & Carmichael, G. G. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4835–4838.
Messing, J., & Vieira, J. (1982) *Gene 19*, 269–276.
Miller, M. S., Buss, E. G., & Clagett, C. O. (1981) *Biochim. Biophys. Acta 677*, 225–233.
Mount, S. M. (1982) *Nucleic Acids Res. 10*, 459–472.
Nardelli, D., van het Schip, F. D., Gerber-Huber, S., Haefliger, J.-A., Gruber, M., AB, G., & Wahli,W. (1987a) *J. Biol. Chem. 262*, 15377–15385.
Nardelli, D., Gerber-Huber, S., van het Schip, F. D., Gruber, M., AB, G., & Wahli, W. (1987b) *Biochemistry 26*, 6397–6402.
Neufeld, E. F., & Ashwell, G. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W. J., Ed.) pp 241–266, Plenum Press, New York.
Öfverstedt, L.-G., Hammarstrom, K., Balgobin, N., Hjerten, S., Pettersson, U., & Chattopadhyaya, J. (1984) *Biochim. Biophys. Acta 782*, 120–126.
Renugopalakrishnan, V., Horowtiz, P. M., & Glimcher, M. J. (1985) *J. Biol. Chem. 260*, 11406–11413.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schubiger, J.-L., & Wahli, W. (1986) *Nucleic Acids Res. 14*, 8723–8734.
Shainkin, R., & Perlmann, G. E. (1971) *J. Biol. Chem. 246*, 2278–2284.
Small, D., Chou, P. Y., & Fasman, G. D. (1977) *Biochem. Biophys. Res. Commun. 79*, 341–346.
Spieth, J., Denison, K., Zucker, E., & Blumenthal, T. (1985) *Nucleic Acids Res. 13*, 7129–7138.
Stifani, S., George, R., & Schneider, W. J. (1988) *Biochem. J. 250*, 467–475.
Taborsky, G. (1980) *J. Biol. Chem. 255*, 2976–2985.
van het Schip, F. D., Samallo, J., Broos, J., Ophuis, J., Mojet, M., Gruber, M., & AB, G. (1987) *J. Mol. Biol. 196*, 245–260.
Wahli, W., Germond, J.-E., ten Heggeler, B., & May, F. E. B. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6832–6836.
Wallace, R. A. (1985) in *Developmental Biology* (Browder, L. W., Ed.) Vol. I, pp 127–177, Plenum, New York.
Wallace, R. A., & Morgan, J. P. (1986) *Biochem. J. 240*, 871–878.
Wang, S.-Y., & Willams, D. L. (1980) *Biochemistry 19*, 1557–1563.
Wang, S.-Y., Smith, D. E., & Williams, D. L. (1983) *Biochemistry 22*, 6206–6212.
Wicks, R. J., & Clark, R. C. (1987) *Int. J. Biochem. 19*, 33–39.
Wiley, H. S., & Wallace, R. A. (1981) *J. Biol. Chem. 256*, 8626–8634.
Young, R. A., & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194–1198.
Yusko, S., Roth, T. F., & Smith, T. (1981) *Biochem. J. 200*, 43–50.

Evidence That the 3′ End of a tRNA Binds to a Site in the Adenylate Synthesis Domain of an Aminoacyl-tRNA Synthetase†

Kelvin Hill‡ and Paul Schimmel*
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received September 26, 1988; Revised Manuscript Received November 10, 1988

ABSTRACT: Aminoacylation requires that an enzyme-bound aminoacyladenylate is brought proximal to the 3′ end of a specific transfer RNA. In *Escherichia coli* alanyl-tRNA synthetase, the first 368 amino acids encode a domain for adenylate synthesis while sequences on the carboxyl-terminal side of this domain are required for much of the enzyme–$tRNA^{Ala}$ binding energy. The 3′ end of *E. coli* $tRNA^{Ala}$ has been cross-linked to the enzyme, and sequence analysis showed that Lys-73 is the major site of coupling. A mutant enzyme with a Lys-73 → Gln replacement has a 50-fold reduced k_{cat}/K_m (with respect to $tRNA^{Ala}$) for aminoacylation but has a relatively small alteration of its kinetic parameters for ATP and alanine in the adenylate synthesis reaction. The data provide evidence that the 3′ end of $tRNA^{Ala}$ binds to a site in the enzyme domain responsible for adenylate synthesis and that a residue (Lys-73) in this domain is important for a $tRNA^{Ala}$-dependent step that is subsequent to the synthesis of the aminoacyladenylate intermediate.

The complementary use of chemical modification and site-directed mutagenesis has been applied recently in a number of cases to probe structure–activity relationships in enzymes [Huynh et al., 1988; Kato et al., 1988; Nickbarg et al., 1988; reviewed in Profy and Schimmel (1988)]. This approach identifies regions of functional interest in a protein by chemical modification techniques yet overcomes many inherent limitations of interpretation through the introduction of unam-

† This work was supported by Grant GM15539 from the National Institutes of Health.
* Correspondence should be addressed to this author.
‡ Recipient of a postdoctoral fellowship from Loma Linda University.

0006-2960/89/0428-2577$01.50/0