Nucleic Acids Research

The complete nucleotide sequence of the chick a-actin gene and its evolutionary relationship to the actin gene family

James A.Fornwald, Gerald Kuncio, Isaac Peng and Charles P.Ordahl[+]

Department of Anatomy, Temple University School of Medicine, 3400 North Broad Street, Philadelphia, PA 19140, USA

ABSTRACT

The nucleotide sequence of the chick a-actin gene reveals that the gene is comprised of 7 exons separated by six very short intervening sequences (IVS). The first IVS interrupts the 73 nucleotide 5' untranslated segment between nucleotides 61 and 62. The remaining IVS interrupt the translated region at codons 41/42, 150, 204, 267, and 327/328. The 272 nucleotide 3' untranslated segment is not interrupted by IVS. The amino acid sequence derived from the nucleotide sequence is identical to the published sequence for chick a-actin except for the presence of a met-cys dipeptide at the amino-terminus. The IVS positions in the chick a-actin gene are identical to those of the rat a-actin gene. While there is partial coincidence of the IVS positions in the a-actin genes with the vertebrate b-actin genes and 2 sea urchin actin genes, there is no coincidence with actin genes from any other source except soybean where one IVS position is shared. This discordance in IVS positions makes the actin gene family unique among the eucaryotic genes analyzed to date.

INTRODUCTION

Actin, one of the oldest known proteins, is a major component of the cytoskeletons of all eucaryotic cells (1) and is the predominant component of the contractile apparatus of muscle cells (2). Cytoskeletal and muscle actins are similar but not identical. In all, 6 different actin iso-proteins have been identified in birds and mammals (3,4,5). Four of these are *a-actins or muscle actins obtained from 4 different muscle tissue types (skeletal striated, cardiac striated, vascular smooth, and enteric smooth). The a-type actins are very similar to one another, having a total of only 9 variable amino acid positions, 5 of which are the first 5 positions at the N-terminus. Not all 9 potentially variable positions differ between individual a-types; for example, the two smooth muscle actins differ only at positions 1,4, and 5. The two non-a-actins are the cytoskeletal (or cytoplasmic) actins termed b- and g-actin. These differ from each other at positions 2,3,4 and 10 but differ from the a-types at a total of 27 positions.

Actins from diverse organisms are also extremely similar. For example,

chick, bovine and rabbit skeletal muscle actins have an identical amino acid
sequence (3,5) which differs from the yeast actin sequence at only 49 out of
375 positions (6,7). The clustering of amino acid changes between actin
isoproteins, and the conservation between widely diverged organisms
demonstrate that a consensus amino acid sequence for actin is maintained by
strong selection pressure. Indeed, this strong evolutionary constraint is
reflected at the nucleotide sequence level since actin genes from widely
disparate organisms efficiently cross-hybridize.

The number of actin genes in different organisms varies widely as esti-
mated by Southern blot hybridization (8). In chick, 4-7 distinct actin bands
have been observed indicating the presence of 4-7 genes (9), as expected from
the available protein sequence data. In human DNA, however, similar experi-
ments indicate the presence of at least 20 actin genes (10). It is not yet
clear whether all these genes encode different functional actin proteins or
whether they represent multiple copies of isogenes or pseudogenes. Amongst
the non-vertebrates the numbers of actin genes also vary widely: yeast-1 gene
(6,7); Drosophila-6 genes (11); slime mold-17 genes (12); sea urchin-11 genes
(13).

The multiplicity of actin proteins and genes within a single organism may
be explained in two ways. First, the micro-diversification of actin amino
acid sequence may confer functional differences important for the specialized
roles of each actin iso-protein. In vitro assays of the ability of different
iso-actins to polymerize, and to interact with myosin demonstrate quantitative
differences which may indicate functional differences between the various iso-
proteins (1). A second, more speculative, hypothesis would state that each of
the actin isogenes is organized within a different regulatory gene set to
facilitate its tissue-specific co-expression with other genes. In this case
the micro-diversification of actin amino acid sequence may be largely irrele-
vant to the actual function of the protein in vivo. Partial support for this
latter idea comes from the observation that the muscle-specific actins in
Drosophila more closely resemble the vertebrate cytoplasmic actins than the
muscle actins of vertebrates (11). At present, the available data cannot
exclude either explanation.

Analysis of the actin gene family can be used to study tissue-specific
gene regulation as well as the evolutionary processes leading to genic diver-
sification. Towards these ends we have begun to analyze the structural orga-
nization and transcriptional regulation of chick actin genes. Here we report
the complete nucleotide sequence of the a-actin gene expressed in leg muscle

and compare its structure to that of actin genes from a number of diverse
organisms.

METHODS AND MATERIALS

DNA SEQUENCING

Cloned DNA was digested with restriction endonucleases (New England
Biolabs, Bethesda Research Labs) according to the suppliers recommendations
and fractionated by gel electrophoresis. Fragments selected for sequencing
were isolated from the gels (14,15) and labeled either at their 5' termini
using polynucleotide kinase and g-^{32}P ATP (ICN) (14) or at their 3' termini
with a-^{32}P cordycepin (Amersham) using terminal transferase (16). The end-
labeled fragments were then re-digested with appropriate restriction endo-
nucleases and the resulting singly end-labeled fragments re-isolated as above.
Sequencing was conducted using the chemical cleavage method essentially as
described by Maxam and Gilbert (14) and the cleaved fragments fractionated on
thin urea-acrylamide gels (17) which were autoradiographed with x-ray film
(Kodak) in the presence of intensifying screens (Dupont, Cronex) at -70°C.
Based on several sequence determinations of the same region using different
labeling sites, we would predict an overall error rate of less than 1%, simi-
lar to that predicted elsewhere (11). One segment, however, was repeatedly
found to be especially difficult to sequence (positions -75 to -35, Figure 2)
and as a result may have a higher error rate.

mRNA SEQUENCING

Dideoxynucleotide terminated sequencing of mRNA was modified from the pro-
cedure as described by Levy et al. (18). A 116 bp HhaI-RsaI fragment spanning
positions 209-324 (Figure 2) was isolated after fractionation on acrylamide
gels as above. Two picomoles of this fragment were denatured in 80ul 98% for-
mamide by heating to 85°C for 10 min. After denaturation, 100ug poly(A) RNA
from day 18 embryonic chick leg muscle in 20ul 200mM Pipes (pH 6.4), 2.0M
NaCl, 5mM EDTA was added and the mixture incubated at 52°C overnight.
Hybridization was terminated and the nucleic acid precipitated by addition of
70ul 5M NH$_4$ acetate and 500ul ethanol. The nucleic acid precipitate was
collected by centrifugation, redissolved in 300mM Na acetate and again precipi-
tated by addition of 2.5 vol. ethanol. The final ethanol precipitate was
collected, dried under vacuum and redissolved in 11ul water on ice. To ensure
complete dissolution of the RNA, 2ul 10x H buffer (340mM Tris-HCl, pH 8.3,
500mM NaCl, 60mM MgCl$_2$, 50mM dithiothreitol; ref. 19) was added and the solu-
tion heated to 65°C for 10 min. followed by chilling on ice. We found that

this heating step substantially improved the resolution of individual bands over background in the sequence ladder. The dissolved template-primer mixture was added to redissolve 100 picomoles of dried a-^{32}P dCTP (Amersham, 3000 Ci/Mole), followed by addition of 1ul (13 units) reverse transcriptase (J. Beard) and 1ul of 1.25mg/ml Actinomycin-D (Sigma) (final vol. 15ul). 3ul aliquots of this mixture were then added to five 1.5ml Eppendorf tubes each of which contained 2ul unlabeled deoxynucleotide triphosphates and one of the four dideoxynucleotide triphosphate terminators, with the 5th tube containing no chain terminators. The final uM concentrations of nucleotides in each sequence reaction tube were:

G reaction-dGTP, 35; dATP,350; dTTP,350; dCTP,2; ddGTP, 2;

A reaction-dGTP,350; dATP, 35; dTTP,350; dCTP,2; ddATP,20;

T reaction-dGTP,350; dATP,360; dTTP, 35; dCTP,2; ddTTP,80;

C reaction-dGTP,260; dATP,260; dTTP,260; dCTP,2; ddCTP,10;

0 reaction-same as C reactions minus dideoxynucleotides.

The reactions were mixed and then incubated at 37°C for 15 min after which 1ul of 1.2mM dCTP in 1x H buffer containing 1 unit of reverse transcriptase was added and incubation continued at 37°C for 10 min to extend chains terminated due to the low concentration of dCTP. Synthesis was stopped by addition of 2ul 1M NaOH, 15mM EDTA followed by heating to 65°C for 5 min. One ul 2M Tris-HCl (pH 7.5) and 1ul 1M HCl was added to neutralize, followed by 50ul 2.5M NH$_4$ acetate with 150ug/ml tRNA carrier, the nucleic acid precipitated with ethanol and then fractionated on sequencing gels and autoradiographed.

S-1 MAPPING EXPERIMENTS

For S-1 mapping, hybridization was performed exactly as described above for sequencing of mRNA. After removal of formamide by ethanol precipitation, hybrids were dissolved in 300mM NaCl, 50mM Na acetate buffer (pH 4.5), 1mM ZnCl$_2$, 20ug/ml denatured salmon sperm DNA and digested with 2500 units S-1 nuclease (Miles) for 30 min at 37°C. After concentration by ethanol precipitation the samples were fractionated on sequencing gels using sequencing ladders as size markers and autoradiographed.

RESULTS AND DISCUSSION

We previously reported the isolation and characterization of the chick skeletal a-actin gene from a charon 4A library of the chick genome (20). The clone containing the skeletal a-actin gene was identified among the many actin gene clones by its ability to hybridize with a cDNA probe derived from the 3' untranslated region of the skeletal a-actin mRNA. This region of the mRNA is

not conserved and is therefore specific for this gene. The a-actin gene was
mapped by restriction endonuclease digestion and blot hybridization to a 6.2
kb Hind III fragment which was subsequently subcloned into the Hind III site
of pBR322 (clone designation pG a-actin 1, Figure 1A). The transcriptional
orientation was determined, and the size and position of several intervening
sequences (IVS) estimated by R-loop analysis (20).

In order to precisely determine the structural features of the a-actin
gene we undertook to determine its nucleotide sequence. Figure 1B shows the
strategy employed to determine the sequence of the 2.5 kb segment within which
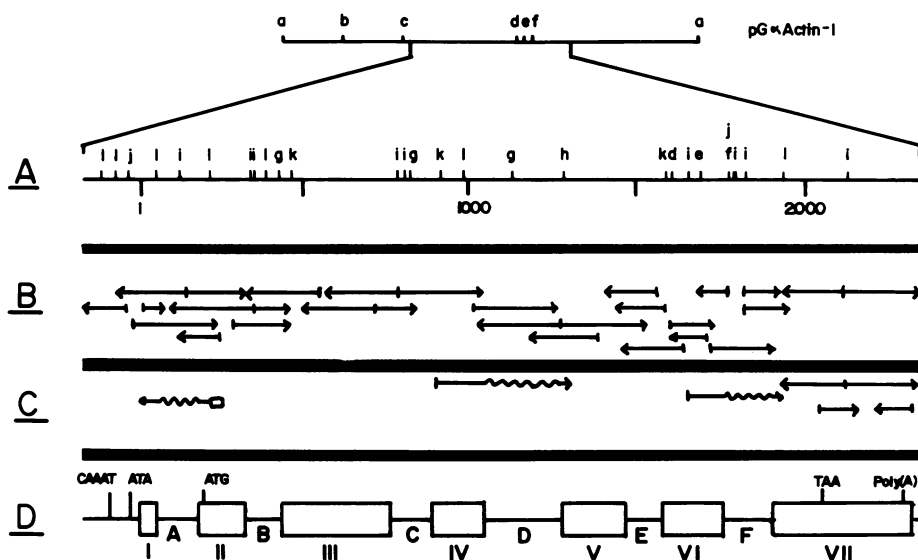the a-actin gene is embedded. In most instances sequencing was performed by

Figure 1. Physical map and sequencing strategy for chick a-actin gene.
Panel A: Upper portion shows the restriction endonuclease cleavage map of the
6.2kb Hind III insert of pGa-actin 1 (20) which contains the a-actin gene.
The lower portion shows a fine restriction map of the 2.5kb segment within
pGa-actin 1 which contains the entire gene. The scale is in nucleotides with
nucleotide 1 being the first nucleotide of the mRNA.
Panel B: Sequencing strategy used to cover the cloned genomic DNA segment.
Panel C: Sequencing performed on cDNA clones and mRNA. Wavy lines indicate
regions where IVS are missing from cDNA clones or mRNA. The run indicated in
the left-hand most region was direct sequencing of the mRNA (see Figure 3).
Panel D: Position of the coding blocks (Roman numerals), IVS (lettered),
transcriptional signals (CAAAT, ATAAA, ATTAAA) and translational signals (ATG,
TAA) in the chick a-actin gene. A summary of the map positions of each exon
and the regions of the mRNA encoded is given in Table 1.
Symbols for restriction endonucleases are; a-Hind III, b-Sac I, c-Sma I, d-Bam
HI, e-Kpn I, f-Xho I, g-Pst I, h-Pvu II, i-Ava II, j-Taq I, k-Hinf I, l-Hha I.

the chemical cleavage method (14) but in one instance the mRNA sequence was
determined using primed cDNA synthesized in the presence of dideoxynucleotide
chain terminators (see below). The sequences of both the genomic segments as
well as relevant cloned cDNA segments were determined (Figure 1, B and C).
Figure 2 shows the nucleotide sequence as determined for the entire 2.5 kb
segment.

REGIONS ENCODING THE SKELETAL a-ACTIN POLYPEPTIDE

The amino acid sequence of chick skeletal a-actin has been previously
determined (3). It was possible, therefore, to compare the known protein
sequence to the translated nucleotide sequence, thereby defining the location
and boundaries of the coding blocks (see Figures 1d and 2). In three instan-
ces the coding block/IVS boundaries could also be deduced from a comparison to
the nucleotide sequence of the mRNA, determined by sequencing cloned cDNAs as
well as direct sequencing of the mRNA (see Figure 1C). The amino acid
sequence for chick skeletal a-actin derived from this translation of the geno-
mic nucleotide sequence is identical to the published sequence (3) except for
the met-cys dipeptide at the N-terminus (Figure 2). Previously, actin was
thought to have a single methionine residue removed prior to acetylation of
the N-terminal aspartate of the mature protein. The data presented here indi-
cate that both a met and a cys residue is removed during this process. The 6
actins of Drosophila are also found to be initiated with a met-cys dipeptide
(11). The derived sequence of the primary a-actin polypeptide is therefore
377 residues. We have numbered the amino acids in Figure 2 in conformity with
the numbering system proposed by Lu and Elzinga (21) which yields 374 posi-
tions. Of the 3 additional positions, 2 are the met-cys dipeptide at the
N-terminus, while the third is the "extra" serine residue between positions
234 and 235 which has been designated position 234a (21).

The precise position of the IVS/exon boundaries cannot always be deter-
mined unequivocally because of nucleotide overlap at the boundaries which has
been observed for several other genes (22,23,24,25). In such cases, the
splice junctions were assigned using the GT/AG rule proposed by Breathnach et
al. (26), which allowed a single assignment of each border. The amino acid

Figure 2. Nucleotide sequence of the chick a-actin gene. The nucleotides are
numbered relative to the putative 1st transcribed nucleotide. Upstream
sequences are given negative numbers relative to this point. Untranslated
regions of the message are underscored; putative control sequences (CAAAT,
ATAAA, see text) are under- and over-scored. Translated nucleotides are
separated into codons with the amino acid specified underneath. Numbering of
the amino acids is as proposed by Lu and Elzinga (21, see also text).

coding region is interrupted by 5 IVS (B-F in Figure 1D) at positions 41/42, 150, 204, 267, and 327/328 (see also Figure 2). The size and positions of the exons and IVS are summarized in Table 1. The position of some of these IVS appear to be unique to vertebrate a-actin genes while others are conserved over great evolutionary distances (see EVOLUTION OF ACTIN GENES, below).

DETERMINATION OF THE 5' EXON BOUNDARIES

While the actin amino acid sequence data and cloned cDNA sequencing could be used to determine the translated regions of the gene, neither was available for a similar determination of the 5' untranslated segment. Using a genomic fragment end-labeled with exon II, we were able to show that this exon extended only 12 nucleotides to the 5' side of the initiation codon, ATG (data not shown). In order to ask whether this represented the 5' end of the mRNA we determined the mRNA 5' end nucleotide sequence directly. A 116 base pair Hha I-Rsa I fragment was isolated which spans positions 209-324 within exon II (Figure 2). This fragment was denatured, hybridized to muscle poly(A) RNA and the template-primer complex used to synthesize cDNAs in the presence of dideoxynucleotide terminators.

Figure 3 shows the results of one such experiment. The primed cDNAs terminate 73 nucleotides upstream from the ATG initiation codon. The sequence matches that of the genomic DNA for the 15 nucleotides upstream from the ATG initiation codon (the last matching nucleotide is designated by an arrow in

Table 1. SUMMARY OF a-ACTIN GENE ORGANIZATION

| Exon/IVS	Positions	size	encoded region of mRNA
Exon I	1-61	61	61 of 73 nt 5' untrans.
IVS A	62-181	111	----
Exon II	182-322	141	aa 1-41 + 12 nt of 5' ut.
IVS B	323-434	112	----
Exon III	435-759	325	aa 42-150
IVS C	760-893	134	----
Exon IV	894-1055	162	aa 150-204
IVS D	1056-1280	225	----
Exon V	1281-1472	192	aa 204-267
IVS E	1473-1577	105	----
Exon VI	1578-1759	182	aa 267-327
IVS F	1760-1908	149	----
Exon VII	1909-2324	416	aa 328-374 + 272 nt 3' ut.
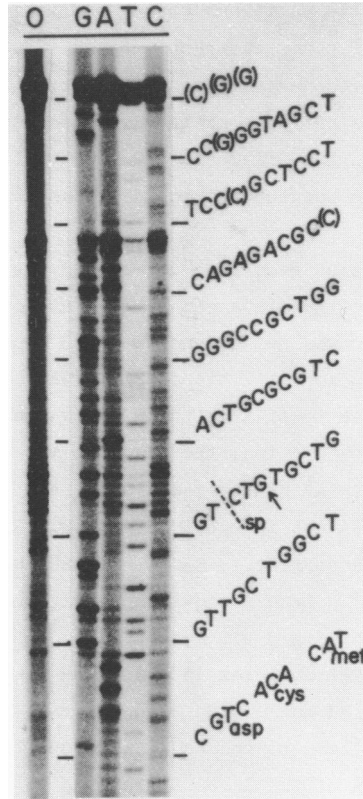
Figure 3. Autoradiograph of sequencing gel showing chain terminated reverse transcription of the 5' and of a-actin mRNA. cDNA synthesis was primed using a genomic fragment derived from exon II (see text and Methods and Materials). Nucleotides complementary to those specifying amino acids are separated into triplets with the specified amino acid shown. The methionine specified is the initiator. The arrow points to the last nucleotide which matches contiguously with exon II (nucleotide 179 in Figure 2). Dashed line labeled "sp" shows the presumptive splice point based upon GT/AG rule (see text). The synthesis lane in which no dideoxynucleotides were added shows that many "strong stops" prematurely terminate synthesis and serve to confuse reading of the sequence. In some cases the residue can be read despite the strong stop, while in others it cannot. Nucleotides which were ambiguous on the original film of this gel are shown in parentheses. For these, the assignment of the correct nucleotide was made from other experiments and/or from the sequence of the genomic fragment.

Figure 3) after which no contiguous match is seen. A match is found in the genomic sequence 114 nucleotides upstream, beginning at position 58 and continuing 58 nucleotides to position 1. This most extreme terminator position is 27 nucleotides downstream from the "T" of an ATAAA sequence and 85 nucleotides downstream from the "C" of a CAAAT sequence (Figure 2). These

sequences and their relative positions correspond well with similar sequences found at the 5' termini of other eucaryotic genes (27,28). We tentatively conclude, therefore, that the extreme termination point of cDNA synthesis represents the 5' terminus of the a-actin mRNA and the likely transcriptional initiation site. Based on the GT/AG rule (26) we would place the junction of exon I and II as occurring between nucleotide 61 and 62 (designated "sp" in Figure 3).

To confirm the borders of exon I we determined its boundaries by S-1 mapping (29) using genomic fragments end-labeled at positions within exon I (Figure 4). The 3' border was defined using an 88 base pair Hpa II fragment spanning the 3' border of exon I. This fragment was labeled at its 3' ends using reverse transcriptase to fill in the 3' recessed end with a-^{32}P dCTP and unlabeled dGTP. After denaturation, the fragment was allowed to hybridize with either leg muscle poly(A) RNA or yeast tRNA, and then digested with S-1 nuclease to degrade the DNA not in hybrid register. If the 3' border of exon I defined by sequencing is correct, a 53 nt end-labeled fragment should survive digestion (Figure 4A). Lane A shows that only the full-length fragment is observed when it is hybridized with yeast tRNA. This full-length fragment represents renatured DNA-DNA duplexes in which the labeled ends are S-1 resistant due to the labeled nucleotide being in base-pair register. When hybridized to muscle RNA (lane b) only a small proportion of renatured fragment is seen while the majority of the label migrates at 4 positions between 51 and 54 nucleotides in length, agreeing well with the predicted length of 53 nucleotides. The variability in the lengths of these digestion products is presumably a result of "nibbling" at the base-paired ends by the S-1 nuclease.

The 5' border of exon I was defined by a similar procedure using a 72 base pair Taq I-Hha I fragment which spans the 5' border (Figure 4B) labeled at its 5' termini using g-^{32}P ATP and polynucleotide kinase. As shown in Figure 4B, a protected fragment 47 nucleotides in length is expected and fragments 45-47 nucleotides are seen after hybridization with muscle RNA but not with yeast tRNA. Little or no full-length fragment is seen in either lane in this case because in the renatured fragment the label at the Taq I site is completely vulnerable to digestion, and at the Hha I site the labeled phosphate in the 5' position is probably more susceptible to being removed by "nibbling" than is the case with the Hpa II fragment where the labeled nucleotide was located one nucleotide in from the extreme 3' end. The results of the S-1 mapping indicate that the borders of exon I as deduced from sequence analysis are correct to within +/- 2 nucleotides. Greater precision is precluded by the variabi-
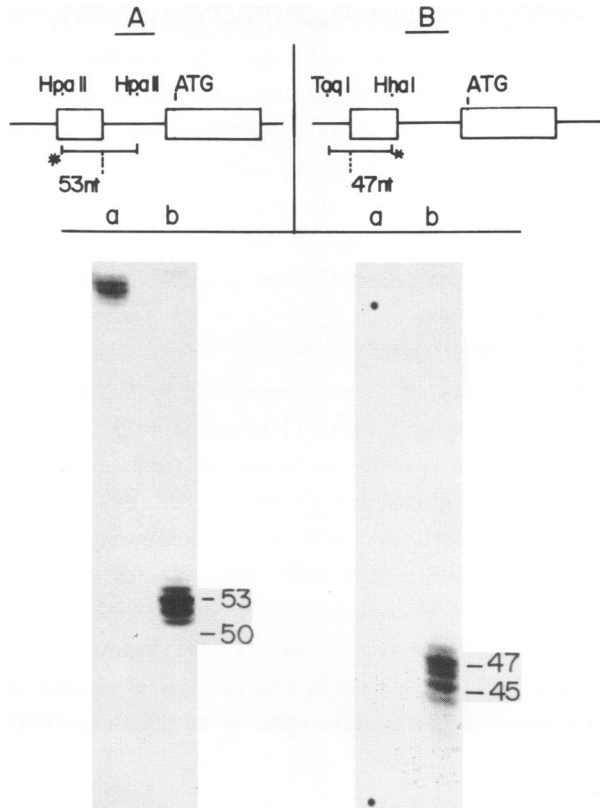
Figure 4. S-1 mapping of exon I. Panel A shows the experiment used to define the 3' border of exon II; Panel B shows that for the 5' border. For each, the drawing at the top shows the labeled fragment used relative to exons I and II. The position of the ATG initiator in exon II is indicated. The dashed line shows the expected position for digestion by S-1 nuclease with the number of nucleotides from that position to the labeled end. Lanes A and B show the digestion products obtained after hybridization of the probe with yeast tRNA or day 18 embryo leg muscle poly(A) RNA, respectively.

lity in length of the end-labeled fragments surviving S-1 nuclease digestion. As a result of this, and the three very strong terminators seen at the 5' end of the mRNA during sequencing (Figure 3), we cannot rule out the possibility that some a-actin mRNAs initiate at variable points between positions -1 to +3.

SEQUENCE OF THE 3' UNTRANSLATED REGION

The cloned cDNAs in our possession, as a group, encoded regions towards the 3' end of a-actin mRNA. Two independent cDNA clones, pa-actin 1 and pa-actin 2 (20), contained complete and intact 3' untranslated regions. The nucleotide sequence of both of these was found to be identical from the TAA

terminator codon through 272 nucleotides to the poly(A) tail. In addition, that sequence was also found to be identical to the corresponding region in the genomic DNA. We conclude, therefore, that no IVS interrupts the 3' untranslated segment of the chick skeletal a-actin gene. The sequence ATTAAA was found 31 nucleotides upstream from the putative poly(A) addition site consistent with the proposal (27) that this sequence may be the poly(A) addition signal. The presence of 2 A residues following the poly(A) addition site designated in Figure 2 make it possible that polyadenylation occurs at any one, or all, of these residues.

SUMMARY OF a-ACTIN GENE STRUCTURE

The a-actin gene consists of 7 exons separated by 6 IVS, the longest of which is IVS D (Figure 1D, 2, and Table 1) which is only 225 nucleotides in length. The 5' untranslated segment is 73 nucleotides in length and is interrupted between nucleotides 61 and 62 by an IVS. The 3' untranslated segment is 272 nucleotides in length and is uninterrupted by IVS. This gives an overall mRNA length of 1476 nucleotides, exclusive of polyadenylation. Consistent with this, recent measurements (not shown) demonstrate that leg muscle a-actin mRNA co-migrates with E. coli 16s ribosomal RNA which is 1541 nucleotides in length (30). Using chick 18s and 28s ribosomal RNA as markers, we had previously estimated the a-actin mRNA to be 2000 nucleotides in length (20).

EVOLUTION OF ACTIN GENES

In the eucaryotic gene families studied to date, the positions of IVS appear to be strongly conserved (25,31,32) even over large evolutionary distances. For example, the globin gene IVS position in amphibians and mammals are identical even though these groups diverged over 400 million years ago (32). This conservation leads to the conclusion that either the selection pressure to maintain IVS position is very strong or that their removal is a relatively rare event. The function, if any, of IVS is not well understood. One hypothesis states that IVS are remnants of the initial events in the evolution of genes which involves the collection of exons from different sources via recombination (33,34). A corollary to this hypothesis states that descendants of an ancestral gene will have either the same IVS positions, or a subset thereof, assuming that IVS removal occurs given sufficiently long time periods.

The actin gene family is unique among the gene families studied to date in that its IVS positions are not conserved between different members. Figure 5 shows the IVS positions for actin genes from a number of different eucaryotes.

IVS Position

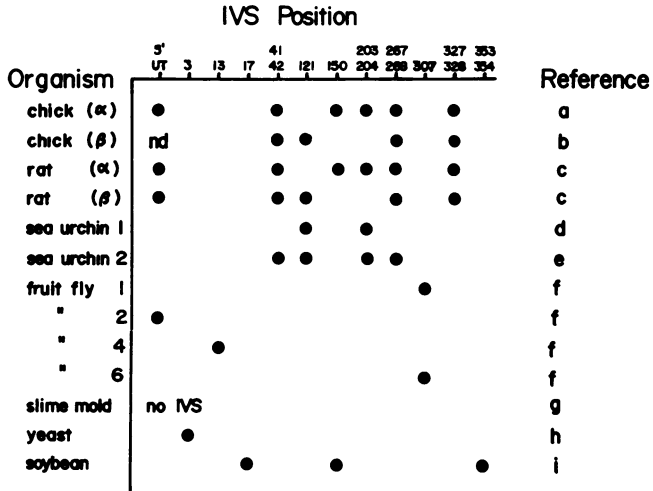| Organism	5' UT	3	13	17	41 42	121	150	203 204	267 288	307	327 328	353 354	Reference
chick (α)	●				●			●	● ●		●		a
chick (β)	nd				●	●			●		●		b
rat (α)	●				●			●	● ●		●		c
rat (β)	●				●	●			●		●		c
sea urchin 1					●		●						d
sea urchin 2					●	●		●	●				e
fruit fly 1										●			f
fruit fly 2	●												f
fruit fly 4			●										f
fruit fly 6										●			f
slime mold	no IVS												g
yeast	●												h
soybean				●			●					●	i

Figure 5. IVS positions in actin genes from various sources. Codon positions are shown across the top. These positions have been numbered relative to the vertebrate sequence and for those actins with different numbering systems (eg. soybean) the vertebrate equivalent position is shown. 5' UT refers to the 5' untranslated region of the message. ND indicates "no data". The references are: a–this report. b–T. Kost and S. Hughes, personal communication; B. Paterson, personal communication. c–U. Nudel et al., in press; D. Yaffe, personal communication. d–A. Cooper and W. Crain, submitted for publication. e–W. Crain and A. Cooper, personal communication. f–reference 11. g–reference 12. h–references 6 and 7. i–reference 40.

Of the 11 different IVS positions identified to date only a few are shared between different members of the actin gene family. The most notable shared IVS positions are found for the a- and b-isogenes from chick and rat. The IVS positions in the a–actin genes of rat and chick are identical, as are those in the b–actin genes of chick and rat. Thus, the a–actin genes and the b–actin genes show a degree of IVS position conservation comparable to that seen between the globin genes of birds and mammals (35). However, only partial homology is seen when the b–actin genes are compared to the a–actin genes because each has IVS at positions where none is found in the other. This suggests, therefore, that these isogenes probably shared a common ancestor prior to the divergence of birds and mammals or more than 300 million years ago (36).

When the chick and rat actin genes are compared to those of a closely related non–vertebrate, the sea urchin, partial, but not complete, homology is again seen. Each of the IVS positions found in the 2 sea urchin genes studied so far has a counterpart in one or both vertebrate genes. Both sea urchin genes, however, contain at least one IVS at a position not found in one or the

other vertebrate gene. Since echinoderms and vertebrates (deuterostomes) shared a common ancestor in the pre-cambrian period of evolution (36) it would appear that a common actin gene IVS pattern established in such an ancestor has been modulated during subsequent evolution.

Essentially no common pattern can be discerned when deuterostomes are com-pared to protostomes (fruit fly) or to lower eucaryotes such as slime mold or yeast, because there is not a single coincident IVS position. Interestingly, one of the 3 IVS positions found in soybean actin gene (at position 150) pre-cisely corresponds to one found in vertebrate a-actin genes. It is also interesting to point out, in this regard, that in the soybean leghemoglobin gene 2 of the 3 IVS positions correspond to the 2 IVS positions in vertebrate globin genes (37).

How such a discordant pattern of IVS positions arose within the actin gene family remains speculative. Others (11) have suggested that this complex pat-tern may have arisen from either insertion of IVS during evolution or from deletion of IVS from a more complex ancestral pattern. While there is no direct evidence for the former, it is clear that IVS deletion can occur, apparently concommitant with chromosomal translocation (25,37,38).

The underlying assumption for the foregoing discussion is that all actin genes share a common, albeit evolutionarily ancient, ancestor. At present this assumption is not verifiable. We cannot, therefore, rule out the possi-bility that the actin genes of deuterostomes and, for example, protostomes might not share a common ancestor. If such is the case, then the similarity in actin amino acid sequence and actin gene nucleotide sequence between insects and vertebrates arose via convergent evolution. Whether mechanisms for convergent evolution at the molecular level exist remains to be determined.

*Abbreviations used: bp-base pair, IVS-intervening sequence; kb-kilobase pair; nt-nucleotide, a-actin, alpha actin; b-actin, beta actin; g-actin, gamma actin.

+To whom correspondence should be sent at the following address: Department of Anatomy, School of Medicine, University of California at San Francisco, San Francisco, CA 94143, USA.

REFERENCES
 1. Korn, E.D. (1978) Proc. Nat. Acad. Sci. USA 75:588-599.
 2. Potter, J.D. (1974) Arch. Biochem. Biophys. 162:436-441.

3. Vanderkerckhove, J. and Weber, K. (1979) FEBS Letters 102:219-222.
4. Vanderkerckhove, J. and Weber, K. (1978) J. Mol. Biol. 126:783-802.
5. Vanderkerckhove, J. and Weber, K. (1979) Differentiation 14:123-133.
6. Gallwitz, D. and Sures, I. (1980) Proc. Nat. Acad. Sci. USA 77:2546-2550.
7. Ng, R. and Abelson, J. (1980) Proc. Nat. Acad. Sci. USA 77:3912-3916.
8. Southern (1975) J. Mol. Bio. 98:503-517.
9. Cleveland, D.W., Lopata, M.A., MacDonald, R.J., Cowan, N.J., Rutter, W.J. and Kirschner, M.W. (1980) Cell 20:95-105.
10. Humphries, S.E., Whittall, R., Minty, A., Buckingham, M. and Williamson, R. (1981) Nucleic Acids Res. 9:4895-4908.
11. Fyrberg, E.A., Bond, B.J., Hershey, N.D., Mixter, K.S. and Davidson, N. (1981) Cell 24:107-116.
12. McKeown, M. and Firtel, R.A. (1981) Cell 24:799-807.
13. Scheller, R.H., McAllister, L.B., Crain Jr., W.R., Durica, D.S., Posakony, J.W., Thomas, T.L., Britten, R.J. and Davidson, E.H. (1981) Mol. and Cell. Biol. 1:609-628.
14. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymolology 65:499-560.
15. Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981) Anal. Biochem. 112:295-298.
16. Tu, C.D. and Cohen, S.N. (1980) Gene 10:177-183.
17. Sanger, F. and Coulson, A.R. (1978) FEBS Letters 87:107-110.
18. Levy, S., Sures, I. and Kedes, L.H. (1979) Nature 279:737-739.
19. Air, G.M., Sanger, F. and Coulson, A.R. (1976) J. Mol. Biol. 108:519-533.
20. Ordahl, C.P., Tilghman, S.M., Ovitt, C., Fornwald, J. and Largen, M.T. (1980) Nucleic Acids Res. 8:4989-5005.
21. Lu, R.C. and Elzinga, M. (1977) Biochem. 16:5801-5806.
22. Jensen, E.O., Paludan, K., Hyldig-Nielsen, J.J., Jorgensen, P. and Marcker, K.A. (1981) Nature 291:677-679.
23. Stein, J.P., Catterall, J.F., Kisto, P., Means, A.R. and O'Malley, B.W. (1980) Cell 21:681-687.
24. Yamada, Y., Avvedimento, V.E., Mudryj, M., Ohkubo, H., Vogeli, G., Irani, M., Pastan, I. and Crombrugghe, B. (1980) Cell 22:887-892.
25. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980) Cell 20:555-566.
26. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) Proc. Nat. Acad. Sci. USA 75:4853-4857.
27. Benoist, C., O'Hare, K. and Chambon, P. (1980) Nucleic Acids Res. 8:127-142.
28. Degrave, W., Derynck, R., Tavernier, J., Haegeman, G. and Fiers, W. (1981) Gene 14:137-143.
29. Weaver, R.F. and Weissman, C. (1979) Nucleic Acids Res. 7:1175-1193.
30. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. USA 75:4801-4805.
31. Eifferman, F.A., Young, P.R., Scott, R.W. and Tilghman, S.M. (1981) Nature 294:713-718.
32. Patient, R.K., Elington, J.A., Kay, R.M. and Williams J.G. (1980) Cell 21:565-573.
33. Gilbert, W. (1978) Nature 271:501.
34. Darnell, J.E. (1978) Science 202:1257-1259.
35. Dodgson, J.B., Strommer, J. and Engel. J.D. (1979) Cell 17:879-887.
36. Waterman, A.J. (1971) Chordate Structure and Function, pp. 42-90. Macmillan, New York.
37. Jensen, E.O., Paludan, K., Hyldig-Nielsen, J.J., Jorgensen, P. and Marcker, K.A. (1981) Nature 291:677-679.
38. Nishioka, Y., Leder, A. and Leder, P. (1980) Proc. Nat. Acad. Sci. USA 77:2806-2809.

39. Hollis, G.F., Hieter, P.A., McBride, O.W., Swan, D. and Leder, P. (1982) Nature 296:321-325.
40. Shah, D.M., Hightower, R.C. and Meagher, R.B. (1982) Proc. Nat. Acad. Sci. USA 79:1022-1026.