Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 10782–10786, November 1992
Evolution

Evolutionary conservation pattern of zinc-finger domains of *Drosophila* segmentation genes

(zinc finger proteins/evolution)

RALF J. SOMMER*, MONIKA RETZLAFF*, KAI GOERLICH*, KLAUS SANDER†, AND DIETHARD TAUTZ*

*Zoologisches Institut der Universität München, Luisenstrasse 14, D-8000 München 2, Federal Republic of Germany; and †Zoologisches Institut der Universität Freiburg, Albertstrasse 21a, D-7800 Freiburg im Breisgau, Federal Republic of Germany

Communicated by Wolfgang Beermann, July 13, 1992

ABSTRACT **A number of genes of the developmental gene hierarchy in *Drosophila* encode transcription factors containing Cys_2His_2 zinc finger domains as DNA-binding motifs. To learn more about the evolution of these genes, it is necessary to clone the homologs, or more correctly the orthologs, from different species. Using PCR, we were able to obtain apparently orthologous fragments of hunchback (*hb*), Krüppel (*Kr*), and snail (*sna*) from a variety of arthropods and partly also from other animal phyla. Sequence alignments of these fragments show that the amino acid differences can normally not be correlated with the evolutionary distances of the respective species. This is due to an apparent saturation of potential replacements within the finger domains, which is also evident from the frequent occurrence of convergent replacements. Another recurrent feature of these alignments is that those amino acids that are directly involved in determining the DNA-binding specificity of the fingers are most conserved. Using *in vitro* bandshift experiments we can indeed show that the binding specificity of a hunchback finger fragment from different species is not changed. This implies that there is a high selective pressure to maintain the regulatory target elements of these genes during evolution.**

Systematic analysis of the genes constituting the developmental genetic hierarchy in *Drosophila* has brought much insight into the molecular details of the early development of this organism (1, 2). This forms a sound basis for studying the evolution of this process by comparing the expression of the respective genes in closely and distantly related species. However, a prerequisite for such a study is that the orthologs of the respective genes can be identified and cloned (we use the distinction ortholog and paralog according to the definition by Fitch (3), with orthologous genes being those with an assumed common descent, while duplicated variants of a gene within a genome are called paralogous). We have focused here on genes containing Cys_2His_2 zinc fingers as the DNA-binding motif. A variety of transcription factors contain these motifs (reviewed in ref. 4), which are characterized by the invariant spacing of two cysteine and two histidine residues, as well as by the occurrence of aromatic amino acids and a leucine at defined positions within the finger structure (5, 6). A number of genes involved in early pattern formation in *Drosophila* belong to this class. Among these are the gap genes hunchback (*hb*) (7) and Krüppel (*Kr*) (8), as well as snail (*sna*) (9), a gene involved in dorsoventral pattern formation. Krüppel and snail are further characterized by the presence of the so-called "H-C link," a conserved stretch of seven amino acids that links the finger domains (10). It is possible to clone new zinc-finger genes by low-stringency hybridization with the finger-domain of Krüppel (10, 11). However, for these genes, the region of clear similarity is usually restricted to the H-C link, while the fingers themselves are fairly diverged. Accordingly, it was so far not possible to clone a true Krüppel orthologous gene from distantly related species. This is different for snail, where an apparent ortholog could be cloned from *Xenopus* (12) by low-stringency hybridization.

Here we have used the PCR process (13) to search for true orthologs of these three genes in a variety of species. We show that it is possible to clone appropriate gene fragments from species belonging to different phyla. The sequences of these fragments‡ show several interesting features, including an indication that those amino acids of the fingers that are directly involved in DNA binding (14) are usually the most conserved ones. This feature provides a useful rationale for the design of PCR primers for the cloning of further zinc-finger orthologs.

To test the inference of a conservation of the binding specificity, we have devised an *in vitro* band-shifting assay that allows analysis of the binding characteristics of a single finger domain. Applying this assay to a hunchback finger fragment from different species shows that the binding specificity is indeed conserved, even though a number of amino acids are diverged.

MATERIALS AND METHODS

Species. The following species were used: *Drosophila melanogaster* (Diptera, Insecta), *Musca domestica* (Diptera, Insecta), *Calliphora vicina* (Diptera, Insecta), *Sciara coprophila* (Diptera, Insecta), *Psychoda cinerea* (Diptera, Insecta), *Apis mellifera* (Hymenoptera, Insecta), *Tribolium castaneum* (Coleoptera, Insecta), *Euscelis plebejus* (Hemiptera, Insecta), *Schulthesia lampyridiformis* (Blattoidea, Insecta), *Locusta migratoria* (Caelifera, Insecta), *Lithobius forficatus* (Chilopoda), *Cryptops apomalans* (Chilopoda), *Artemia salina* (Branchiopoda, Crustacea), *Pholcus phalangoides* (Arachnida, Chelicerata), *Cupiennius salei* (Arachnida, Chelicerata), *Bithynia tentaculata* (Gastropoda, Mollusca), *Platynereis dumerilii* (Polychaeta, Annelida), and *Oryzias latipes* (Pisces, Vertebrata).

Primer. The PCR primers were designed according to the criteria given in ref. 15. The following primers were used (all in 5′-3′ direction): hunchback proximal primer, AARCACCAYYTNGARTAYCA, where R is G or A, Y is T or C, and N is G, A, T, or C; hunchback distal primer 1, ATGCAGYTTSAGCSWRTGRCA, where S is G or C and W is A or T; and hunchback distal primer 2, GTGWGMRTAYTTRCKCARRTG, where M = A or C, and K = T or G (positions 5524–5543, 5713–5692, and 5730–5710 in the hunchback sequence given in ref. 7); Krüppel proximal primer 1, TAYAARCAYGTGYTRCARAAYCA; Krüppel proximal primer 2, GATCATCAYYTSAARACNCA; and Krüppel distal primer, YTTYARYTGRTTRSWRTCRS-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L01587–L01617).

WRAA (positions 1069–1092, 1156–1173, and 1336–1312 in the Krüppel sequence given in ref. 8); and snail proximal primer 1, ATGGGMYTRWSYAARCA; snail proximal primer 2, GGAGCAYTGAARATGCA; and snail distal primer, TGWGCTCKCAGATTASWNCKRTC (positions 935–952, 1040–1057, and 1220–1196 in the snail sequence given in ref. 9).

PCR and Cloning. PCR (13), cloning, and sequencing of the fragments was performed as described (16). In short, primary PCR reactions were blotted and hybridized under medium stringency with the respective *Drosophila* finger domains. The appropriate size region was then isolated from an agarose gel and reamplified. These fragments were blunt-end-cloned into a phage M13 vector, and the resulting plaques were again hybridized with the *Drosophila* probe. Most cloned fragments were then hybridized under high stringency to a genome blot of the respective species to validate their origin. The fragments that could not be tested in this way are indicated in Fig. 1.

Band Shift Assay. hunchback finger fragments representing positions 1–62 in Fig. 1 from the different species were amplified by PCR, cloned into the *Sma* I site of M13mp19 and sequenced to check their integrity and the orientation. They were then cut out with *Bam*HI/*Bgl* II and cloned into the *Bam*HI site of the pET3a vector (17). The full first finger domain of hunchback was blunt-end-cloned as a *Dde* I fragment (positions 5244–5910 in the sequence given in ref. 7) into the *Bam*HI site of pET3a. The proteins were expressed in BL21 cells (17) by induction with 4 mM IPTG (17). The protein extracts were prepared as described in (18). The finger peptide fragments were identified with a polyclonal antibody produced against the synthetic peptide NH_2-Cys-Val-Asn-Lys-Ser-Met-Leu-Asn-Ser-COOH (positions 26–34 in Fig. 1). The relative amounts of protein used were estimated from titration series on Western blots (see text). Band shifts were done with a double-stranded synthetic oligonucleotide 5′-AGATCTATCAAAAAAATGGATCC-3′ containing a natural hunchback binding site from the Krüppel gene (19). The band-shift assays (10 μl) included 100 pg of ^{32}P-end-labeled oligonucleotide, 100 ng of poly(dI·dC), 10 ng of poly(dA·dT), 100 ng of salmon sperm DNA, 0.1% Nonidet P-40 in 10 mM sodium phosphate buffer (pH 7.7). The protein extracts (10 μl) were appropriately diluted in buffer B (18) and mixed with the above buffer on ice. After incubation for 10 min, the samples were electrophoresed in a 4% polyacrylamide gel in 0.25× TBE (1× TBE 89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.2) for 4 hr at 10 V/cm. The competition samples were first mixed with the competing extract, incubated for 10 min, and then mixed with the extract containing the full hunchback finger domain.

RESULTS

Cloning of hunchback. hunchback belongs to the class of gap genes in *Drosophila* and plays a crucial role in the earliest pattern formation decisions by integrating the function of the anterior and the posterior maternal systems (20, 21). The gene codes for a zinc-finger protein with two separate finger domains, the first with four fingers and the second with two

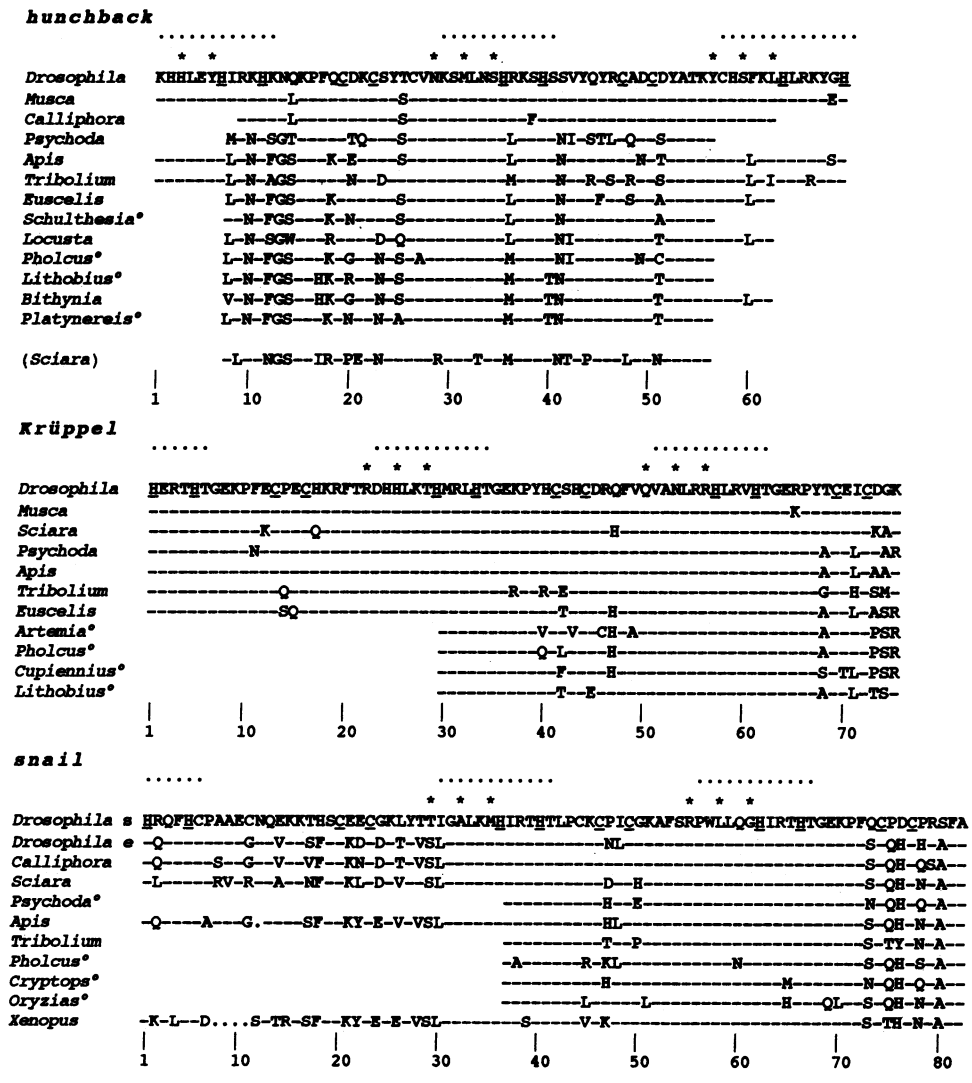

FIG. 1. Sequence alignments of the finger fragments obtained from the genes of different species. (*Top*) hunchback. (*Middle*) Krüppel. (*Bottom*) snail. The sequences are aligned with respect to the D. melanogaster sequence. Identities are denoted by dashes and deletions by dots. Note that different sets of primers have been used for the different species, resulting in differences in the fragment sizes. The row of dots above the sequences indicates the position of the α-helices; the stars denote those amino acids that are believed to be in direct contact with the bases in the DNA (14). The fragments of the species denoted with a circle have not yet been tested in genome blotting experiments. In the hunchback comparisons, we have also included a fragment that was recovered from the lower dipteran *Sciara*, but which we do not consider to be a true ortholog (see text). In the snail gene comparisons, we have included also the sequence of the *Drosophila* escargot gene (ref. 25; see text), denoted by "Drosophila e" and the sequence of the presumptive snail ortholog from *Xenopus* (12).

fingers (7). We have focused our work on the first domain. Two different sets of PCR primers were used for the cloning of hunchback orthologs. One primer pair spans the region between the second and the fourth finger, whereby the 3′ primer lies in the region between the two last structural histidine residues. The other pair uses the same 5′ primer and a 3′ primer that lies in the region preceding the histidine residues of the last finger (see below). These primer pairs allowed the cloning of hunchback from a variety of different species (Fig. 1 *Top*). The sequences suggest clearly that the fragments are true orthologs. hunchback shows a unique feature among the zinc-finger proteins—namely, the absence of an otherwise conserved aromatic amino acid within each finger (positions 26 and 54 (Fig. 1 *Top*) should normally be phenylalanine). However, this may be structurally compensated by the tyrosine residues at positions 24 and 52 (22). The cloned fragments show the same feature—namely, absence of the aromatic amino acid and the presence of the tyrosine at the respective positions. Further proof that the true orthologs of hunchback were recovered comes from the cloning and sequencing of genomic or cDNA clones by using the respective cloned PCR fragments as probes, which showed further characteristic sequence similarities (not shown).

The region that is most conserved in the comparison between the PCR-fragment sequences are the amino acids preceding the structural histidine residue in each finger loop (Fig. 1 *Top*). This region forms part of an α-helix (22, 23) that is directly involved in DNA binding (14). On the basis of this observation, we have designed further PCR primers that lie within this highly conserved region. Utilizing these allowed us to clone hunchback orthologs from species that could not be obtained with previous primer pairs. hunchback orthologs have thus been recovered from a representative selection of arthropods and from two other phyla—namely, a mollusc and an annelid (Fig. 1 *Top*).

hunchback Sequence Conservation Pattern. A number of inferences can be drawn from the pattern of amino acid replacements seen in the hunchback fragments. There is an apparent saturation of possible amino acid replacements, which become evident in the comparisons of the species that have been separated for 200 Myr or more. The sequence difference, for example, between *Drosophila* and *Apis* is not higher than that between *Drosophila* and the mollusc *Bithynia* (24% vs. 28%; 200 Myr vs. 550 Myr). Furthermore, the distance matrix (Table 1) indicates that some of the most distantly related species show a lower divergence than some of the more closely related ones. One interpretation for this effect is that only certain amino acids are allowed at certain positions, which results in convergent replacements. Such apparent convergences can indeed be seen at multiple positions (e.g., positions 20, 23, 36, and 51). Nonetheless, even when taking this effect into consideration, it appears that the fragments obtained from the dipterans are on average more distant to those of the other species than the fragments of these are among themselves (Table 1). The average pairwise distance between the nondipterans is 17%, while the average distance between the higher dipterans and the nondipterans is 26%, with the lower Dipteran *Psychoda* being equally distant to both groups (32% and 29%). This suggests a higher rate of evolution in the dipterans, though additional data will be necessary to show this more conclusively.

Cloning of Krüppel. Employing the rationale to place the PCR primers in those regions of the finger that precede the first structural histidine residue allowed us to obtain also orthologs of the Krüppel gene from a variety of arthropods (Fig. 1 *Middle*). The sequence comparisons show less replacements than for hunchback. This is probably due to a high selective pressure acting not only on the finger loops but also on the H-C link regions. Nonetheless, convergent amino acid replacements can again be found (e.g., positions 47 and 71).

We have failed so far to obtain Krüppel orthologs from other phyla than arthropods. This may be due to the fact that in some species the Krüppel gene contains an intron within the finger region that could be too large for a successful PCR amplification. We have found a small intron in *Apis*, interrupting the first finger domain (in the codon of the arginine at position 19). This intron is not present in *Euscelis* but may be present in *Tribolium*, since we could obtain the respective Krüppel fragment only with a primer that is located 3′ to the intron (note that the full Krüppel sequence in Fig. 1 was derived from a cDNA clone). This indicates that the presence or the absence of an intron can not easily be predicted for Krüppel. This may in fact be a general problem for zinc-finger genes. The location of introns within the finger domains of fully sequenced genes can be variable and defies strict rules (24).

Cloning of snail. The PCR primers for snail were designed both on the basis of the above rationale and on the basis of the comparison with the *Xenopus* snail ortholog (12). Thus, it was possible to obtain apparent snail orthologous fragments from a variety of arthropods, though it was again necessary to design an additional primer (Fig. 1 *Bottom*). snail shows an interesting conservation/divergence pattern. The first finger domain is fairly diverged, while the second and third finger

Table 1. Distance matrix for the hunchback sequence comparisons

	Mu	Ca	Ps	Ap	Tr	Eu	Sch	Lo	Ph	Li	Bi	Pl
Drosophila	2	4	16	12	13	12	10	12	15	14	14	13
Musca (Mu)	—	2	15	11	14	11	9	12	14	13	13	13
Calliphora (Ca)		—	16	12	15	12	9	13	14	14	14	14
Psychoda (Ps)			—	13	13	12	12	12	15	16	17	16
Apis (Ap)				—	11	5	4	8	6	6	9	6
Tribolium (Tr)					—	11	10	11	12	11	13	9
Euscelis (Eu)						—	4	9	8	8	11	8
Schulthesia (Sch)							—	9	8	7	7	6
Locusta (Lo)								—	10	10	12	9
Pholcus (Ph)									—	7	9	7
Lithobius (Li)										—	4	3
Bithynia (Bi)											—	6
Platynereis (Pl)												—

For calculating the pairwise distances, we have only evaluated positions 8–56 (Fig. 1), since these are available for all species studied. The straight number of amino acid replacements for each comparison is given. The average percent difference between the nondipterans and the higher dipterans (see text) was calculated by taking the mean of the amino acid differences from the upper and the lower box, respectively.

domains are much more conserved. Most surprisingly, however, it appears that the *Drosophila* snail sequence is the most distant one in all pairwise comparisons. The snail fragments recovered from the other species are in fact much more similar to a recently described paralog of snail in *Drosophila*, the escargot gene (25) (compare line "Drosophila e" in Fig. 1 *Bottom*). Therefore, the sequence comparisons alone would suggest that we have recovered only escargot orthologs and no snail orthologs. On the other hand, we have obtained always only one variant from all species tested, suggesting that these genes are not normally duplicated in other animals, as they are in *Drosophila* (25). Most interestingly, however, the expression pattern of the respective fragment in *Tribolium* resembles only the snail and not the escargot expression in *Drosophila* (R.J.S. and D.T., unpublished data), suggesting that it is the true snail ortholog.

hunchback Binding Specificity. The observation that the amino acids in the α-helix of the finger region are usually conserved between the different orthologs suggests that the binding specificities are also conserved. We have tested this inference for hunchback, where the consensus binding site in *Drosophila* is fairly well defined (19, 26, 27). However, since only parts of the first finger domain were cloned from the different species, we could use only a single full finger for these experiments. Single fingers, on the other hand, do not bind readily to DNA (28); therefore, it was necessary to devise a different test. Our assay is based on the specific inhibition of the binding of the first finger domain from *Drosophila*. This finger domain is used for band-shifting an oligonucleotide with a consensus binding site for hunchback. This band shift can be specifically inhibited with a single finger peptide from hunchback but not from Krüppel (Fig. 2*A*). A further control shows the high specificity of this assay. A single finger peptide with a mutation in an amino acid that can be expected to be involved in the recognition of the bases in the DNA (position 31 in Fig. 1 *Top*) does not compete with the binding (Figure 2*A*). Single fingers from *Drosophila*, *Musca*, *Psychoda*, *Apis*, *Tribolium*, *Euscelis*, *Locusta*, and *Bithynia* were tested in this assay. The amounts of protein extract were titrated such that the range between lack of inhibition and full inhibition was covered (Fig. 2*B*). The same amounts of protein extract were then tested in a Western blot with an antibody that recognizes specifically the α-helix in this finger domain (see *Materials and Methods*). This allows a good quantitation of the amount of finger protein in the extracts (Fig. 2*B*). The result of this experiment shows that the fingers of all eight species tested inhibit the binding of the hunchback domain from *Drosophila* within a very narrow concentration range (±15%). We take this as a proof that the binding specificity of the respective fragments is conserved, even though they show a number of amino acid replacements outside of the α-helix.

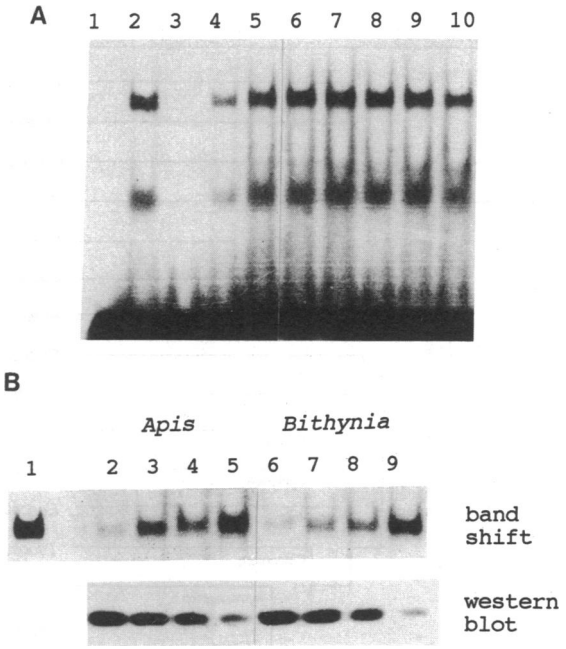

FIG. 2. Band-shift competition experiments with the hunchback finger fragments. (*A*) Test of specificity of competition. Lanes: 1, pET-extract; lanes 2–10, 100 ng of hunchback finger domain extract without (lane 2) or with the following competition extracts: 2.5 (lane 3), 2.0 (lane 4) and 1.5 (lane 5) μg of *Drosophila* hunchback protein (finger fragment); 2.0 (lane 6), 3.0 (lane 7), and 4.0 (lane 8) μg of *Drosophila* hunchback protein (finger fragment) containing a Met-31 → Ile change; 4.0 μg of *Drosophila* Krüppel protein (finger fragment) (lane 9); and 4.0 μg of *Drosophila* Krüppel protein [full-finger domain extract (19)]. Note that a competition is achieved only with the wild-type hunchback finger fragment, but not with the mutated one or with the Krüppel fingers. (*B*) Example of the competition titration tests. (*B Upper*) Bandshifts as in *A*. (*B Lower*) Western blot with the respective amounts of extract. Lanes: 1, no competition; 2–6, competition with decreasing amounts of hunchback protein (finger fragment) from *Apis*; 7–11, competition with decreasing amounts of hunchback protein (finger fragment) from *Bithynia*.

DISCUSSION

We show that it is possible to clone orthologs of *Drosophila* segmentation genes containing a Cys_2His_2 zinc-finger motif by PCR from other species. In our screens, we have usually recovered only one variant of the respective genes, indicating that the three genes under study are not normally members of families of paralogous genes in the species tested. However, we have some indication that Krüppel might be duplicated in the hymenopterans (29). Furthermore, on the basis of the criteria given below, we think that the hunchback sequence recovered from *Sciara* is more likely a paralog rather than an ortholog. Nonetheless, the great majority of the fragments recovered suggests that we are dealing with true orthologs.

The sequence alignments of the different finger fragments show that the amino acids preceding the structural histidine residues are among the most conserved ones. On the one hand, this observation simplifies the design of primers for other cloning projects; on the other hand, it provides also some insight into the function of zinc-finger proteins. The respective amino acids form part of an α-helix that is directly involved in contacting the bases in the DNA and thus in determining the DNA-binding specificity of the protein (14). However, both functional studies (30) and x-ray crystallographic studies (14) have shown that only three amino acids within this region contact the bases directly (indicated in Fig. 1). Accordingly, if there is a strong selection for maintaining this binding specificity, one would expect that only these three amino acids need to be conserved, while the structural α-helix could be formed by a number of different amino acid combinations. The fact that the whole stretch of amino acids is conserved suggests instead that all of them may in some way be involved in the determination of the exact structure of this region. This has partly been observed in the crystal structure, where an amino acid within the helix makes a van der Waals contact with the neighboring amino acid to position it correctly (14). Other interactions could occur via indirect H-bonding through the first shell of water molecules, which would be difficult to detect in crystal structures. These interactions may also be too weak to be detectable by biochemical binding experiments. A role for these amino acids has indeed not been found in such experiments (30, 31), though a function for protein–protein contacts was suggested in one case (31). The fact that this region is evolutionary highly conserved suggests, however, that these amino acids

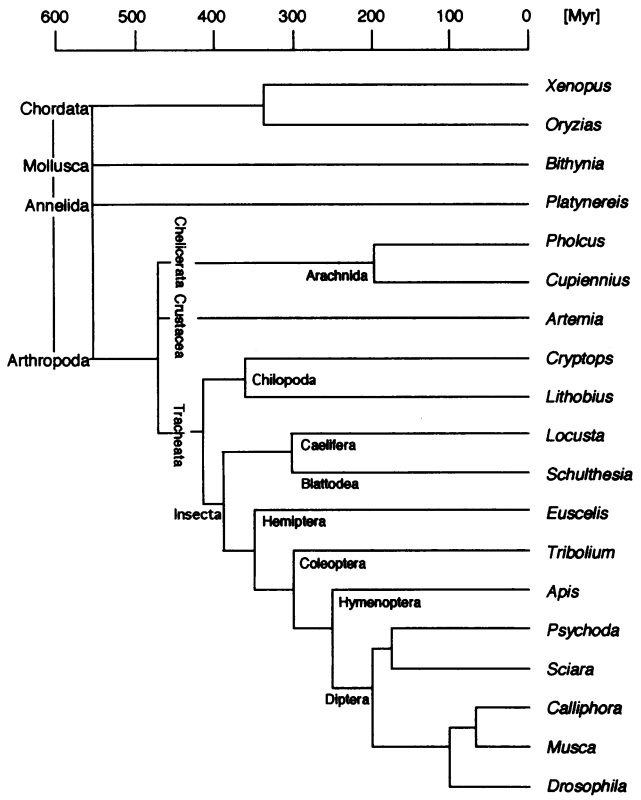

FIG. 3. Phylogenetic tree of the species analyzed in this study. The tree is based on the assumption that all animal phyla have arisen during the Cambrian, 600–550 Myr ago (33). The time scale for the splits of the arthropod lineages has in part been taken from ref. 34 but should be seen only as a rough guideline.

have a general function, most likely in the fine tuning of the structure of the helix and thus in the determination of the DNA binding specificity.

The consideration that the binding specificity of the proteins should be conserved implies that the best criterium for orthology among finger proteins may be the conservation of those amino acids that determine this DNA-binding specificity. These are usually also sufficiently different between different zinc-finger protein genes to make each gene unique. These criteria would fail of course for recently duplicated genes, as appears to be the case for the snail–escargot comparison in *Drosophila*. Nonetheless, given that thousands of zinc-finger genes may exist in the average eukaryotic genome (reviewed in ref. 32), we believe that the criteria given above can be useful.

The general degree of conservation of hunchback in the different species suggests that it should be similarly conserved in deuterostomian lineages, such as echinoderms (e.g., sea urchin) or chordates (e.g., mouse or *Xenopus*). The reason to expect this is that it is now believed that all of the major animal phyla have arisen during a comparatively short time in the Cambrian about 550 Myr ago (Fig. 3) (33). Thus, the time span that separates the origin of the phyla is much smaller then the time of their independent existence. Accordingly, one should expect that genes that are conserved between two different animal phyla should have been present in the common ancestor of all animal phyla, since the time window for the evolution of completely new genes was probably too small. Failure to find a particular gene in other lineages could have two reasons. Either the gene was lost in this particular lineage, or it came under different selective constraints such that too many amino acid positions have changed and the recognizable homology was lost. Our results imply that such selective constraints may exist. We find for the hunchback comparisons that the speed of amino acid replacements may be enhanced in certain lineages, such as the dipterans (Table 1). A similar effect may have caused the otherwise almost unexplainable snail–escargot divergence. Thus, any phylogenetic inferences that are based on such highly constrained DNA-binding domains in proteins should be treated with caution.

We thank T. Gibson for valuable suggestions on a previous version of this manuscript; H. Bohn, W. Büsen, W. Engels, C. Meier-Brook, A. Minelli, and A. Dorresteijn for providing animals used in this study; M. Hoch for providing the oligonucleotide with the hunchback binding site and the Krüppel band-shift extract; M. Hülskamp for providing the information on the mutant hunchback allele; and the members of the laboratory for discussions. This work was supported by D.F.G. Grant Ta99/3-1 to D.T. and by a Ph.D. studentship to R.J.S. by the Fond der chemischen Industrie.

1. Ingham, P. (1988) *Nature (London)* **335**, 25–34.
2. Tautz, D. (1992) in *Development, The Molecular Genetic Approach*, eds. Russo, V. E. A., Brody, S., Cove, D., Ottolenghi, S. & Ruvkun, G. B. (Springer, Heidelberg).
3. Fitch, W. M. (1970) *Syst. Zool.* **19**, 99–113.
4. Johnson, P. F. & McKnight, S. L. (1989) *Annu. Rev. Biochem.* **58**, 799–839.
5. Miller, J., McLaghlan, A. D. & Klug, A. (1985) *EMBO J.* **4**, 1609–1614.
6. Brown, R. S., Sander, C. & Argos, P. (1985) *FEBS Lett.* **186**, 271–274.
7. Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. & Jäckle, H. (1987) *Nature (London)* **327**, 383–389.
8. Rosenberg, U. B., Schröder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. & Jäckle, H. (1986) *Nature (London)* **319**, 336–339.
9. Boulay, J. L., Dennefeld, C. & Alberga, A. (1987) *Nature (London)* **330**, 395–398.
10. Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R. & Jäckle, H. (1986) *Cell* **47**, 1025–1032.
11. Chowdhury, K., Deutsch, U. & Gruss, P. (1987) *Cell* **48**, 771–778.
12. Sargent, M. G. & Bennett, M. F. (1990) *Development* **109**, 967–973.
13. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Ehrlich, H. A. (1988) *Science* **239**, 487–491.
14. Pavletich, N. P. & Pabo, C. O. (1991) *Science* **252**, 809–817.
15. Sommer, R. & Tautz, D. (1989) *Nucleic Acids Res.* **17**, 6749.
16. Sommer, R. & Tautz, D. (1991) *Development* **113**, 419–430.
17. Rosenberg, A. H., Lade, B. N., Chui, D., Lin, S., Dunn, J. J. & Studier, F. W. (1987) *Gene* **56**, 125–135.
18. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987) *Cell* **51**, 1079–1090.
19. Hoch, M., Seifert, E. & Jäckle, H. (1991) *EMBO J.* **10**, 2267–2278.
20. Hülskamp, M., Pfeifle, C. & Tautz, D. (1990) *Nature (London)* **346**, 577–580.
21. Hülskamp, M. & Tautz, D. (1991) *BioEssays* **13**, 261–268.
22. Berg, J. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 99–102.
23. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A. & Wright, P. E. (1989) *Science* **245**, 635–637.
24. Tso, J. Y., Van den Berg, D. J. & Korn, L. J. (1986) *Nucleic Acids Res.* **14**, 2187–2200.
25. Whiteley, M., Noguchi, P. D., Sensabaugh, S. M., Odenwald, W. F. & Kassis, J. A. (1992) *Mech. Dev.* **36**, 117–127.
26. Stanojevic, D., Hoey, T. & Levine, M. (1989) *Nature (London)* **341**, 331–335.
27. Treisman, J. & Desplan, C. (1989) *Nature (London)* **341**, 335–337.
28. Frankel, A. D., Berg, J. M. & Pabo, C. O. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4841–4845.
29. Sommer, R. (1992) Ph.D. thesis (University of Munich).
30. Nardelli, J., Gibson, T. J., Vesque, C. & Charnay, P. (1991) *Nature (London)* **349**, 175–178.
31. Thukral, S. K., Morrison, M. L. & Young, E. T. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9188–9192.
32. El-Baradi, T. & Pieler, T. (1991) *Mech. Dev.* **35**, 155–169.
33. Whittington, H. B. & Morris, S. C. (1985) *Philos. Trans. R. Soc. London Ser. B* **311**, 1–192.
34. Hennig, W. (1969) *Die Stammesgeschichte der Insekten* (Kramer, Frankfurt).