3. J. T. Stull, M. H. Nunnally, C. H. Michnoff, in *The Enzymes: Control by Phosphorylation*, P. D. Boyer and E. G. Krebs, Eds. (Academic Press, New York, 1986), vol 17, p. 113.
4. J. M. Anderson, H. Charbonneau, H. P. Jones, R. O. McCann, M. J. Cormier, *Biochemistry* **19**, 3113 (1980).
5. R. Ranjeva, G. Refeno, A. M. Boudet, D. Marme, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5222 (1983).
6. A. C. Harmon, C. Putnam-Evans, M. J. Cormier, *Plant Physiol.* **83**, 830 (1987).
7. C. Putnam-Evans, A. C. Harmon, M. J. Cormier, *Biochemistry* **29**, 2488 (1990).
8. The lambda bacteriophage cDNA library was obtained from Clonetech, Palo Alto, CA.
9. J. F. Harper, L. Manney, N. D. DeWitt, M. H. Yoo, M. R. Sussman, *J. Biol. Chem.* **265**, 13601 (1990).
10. P. P. Mueller and A. G. Hinnebush, *Cell* **45**, 201 (1986).
11. Y. Nishizuka, *Nature* **334**, 661 (1988).
12. J. Harper, unpublished cDNA sequence from soybean CDPK clone pSK2; *Arabidopsis* cDNA clone pAk1, and genomic clone pgAK1.
13. S. K. Hanks, A. M. Quinn, T. Hunter, *Science* **241**, 42 (1988).
14. M. K. Bennett and M. B. Kennedy, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1794 (1987).
15. T. J. Lukas, D. B. Iverson, M. Schleicher, D. M. Watterson, *Plant Physiol.* **75**, 788 (1984).
16. R. M. Tufty and R. H. Kretsinger, *Science* **187**, 167 (1975).
17. D. M. Roberts, T. J. Lukas, D. M. Watterson, *Crit. Rev. Plant Sci.* **4**, 311 (1986).
18. K. Suzuki, *Trends Biochem. Sci.* **12**, 103 (1987).
19. G. Hardie, *Nature* **335**, 592 (1988).
20. C. Putnam-Evans, A. C. Harmon, B. A. Palevitz, M. Fechheimer, M. J. Cormier, *Cell Motil. Cytoskeleton* **12**, 12 (1989).
21. G. E. Schaller, A. C. Harmon, M. R. Sussman, in preparation.
22. W-7 is (*N*-(6-aminohexyl)-5-chloro-1-napthalene sulfonamide).
23. H. Charbonneau *et al.*, *Biochemistry* **24**, 6762 (1985).
24. M. A. Frohman, M. K. Dush, G. R. Martin, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8998 (1988).
25. We thank S. Kumar and L. Manney for technical assistance with purification of peptides and cDNA sequencing, respectively, M. J. Cormier for fostering the early work on CDPK, R. Amasino for soybean RNA, and K. Walsh for advice and helpful discussions. Supported by grants from the USDA (M.R.S. and A.C.H.), DOE (M.R.S.), and NSF (A.C.H.) and a grant to K. Walsh from NIH (GM15731).

13 November 1990; accepted 7 March 1991

Cloning of a Factor Required for Activity of the Ah (Dioxin) Receptor

EMILY C. HOFFMAN,* HERMINIO REYES,* FONG-FONG CHU,† FRED SANDER,‡ LINDA H. CONLEY, BARBARA A. BROOKS, OLIVER HANKINSON§

The aryl hydrocarbon (Ah) receptor binds various environmental pollutants, such as polycyclic aromatic hydrocarbons, heterocyclic amines, and polychlorinated aromatic compounds (dioxins, dibenzofurans, and biphenyls), and mediates the carcinogenic effects of these agents. The complementary DNA and part of the gene for an 87-kilodalton human protein that is necessary for Ah receptor function have been cloned. The protein is not the ligand-binding subunit of the receptor but is a factor that is required for the ligand-binding subunit to translocate from the cytosol to the nucleus after binding ligand. The requirement for this factor distinguishes the Ah receptor from the glucocorticoid receptor, to which the Ah receptor has been presumed to be similar. Two portions of the 87-kilodalton protein share sequence similarities with two *Drosophila* proteins, Per and Sim. Another segment of the protein shows conformity to the consensus sequence for the basic helix-loop-helix motif found in proteins that bind DNA as homodimers or heterodimers.

THE AH RECEPTOR IS DETECTABLE IN many tissues and organs. The best understood activity of the receptor concerns its role in the induction of cytochrome P450IA1. Complexes between the Ah receptor and ligand bind to specific DNA sequences upstream of the P450IA1 gene, termed xenobiotic responsive elements (XREs), and stimulate transcription of the gene (*1, 2*). The receptor also mediates induction of cytochrome P450IA2 and several other enzymes of xenobiotic metabolism. The enzymatic activities of P450IA1 (aryl hydrocarbon hydroxylase, or AHH) and P450IA2 are important in the metabolism of polycyclic aromatic hydrocarbons (found in cigarette smoke and smog) and certain heterocyclic amines (found in cooked meat) to carcinogenic intermediates (*3*). The pathological effects of the polychlorinated aromatic compounds also depend on the action of the Ah receptor, but the mechanism of pathogenesis is unknown (*4*).

The Ah receptor is a soluble protein complex of ~280 kD. The ~95-kD ligand-binding subunit, which has not been cloned, and the 90-kD heat shock protein (Hsp90) are both components of this complex (*5*). After cells are treated with ligand, receptor molecules become tightly bound in the nucleus. However, the location of the receptor before ligand binding has not been fully resolved (*6*). After conventional subcellular fractionation, the unoccupied receptor is found in the cytosol. Operationally, therefore, ligand treatment leads to "nuclear translocation" of the receptor. The Ah receptor resembles the steroid hormone receptors (7), and this resemblance has led to the suggestion that the ligand-binding subunit of the Ah receptor may be a member of the steroid receptor superfamily.

The mouse hepatoma cell line Hepa-1 shows P450IA1 inducibility. Mutants of Hepa-1 cells defective in induction have been isolated, and those mutants that are recessive have been assigned to four complementation groups (*8, 9*). Mutations in groups B, C, and D affect functioning of the Ah receptor (*9, 10*). In the group C mutants, the receptor is present in normal amounts but does not translocate to the nucleus after binding ligand. We now describe the isolation of part of the human C gene [termed the Ah receptor nuclear translocator gene (*arnt*)] and the isolation and

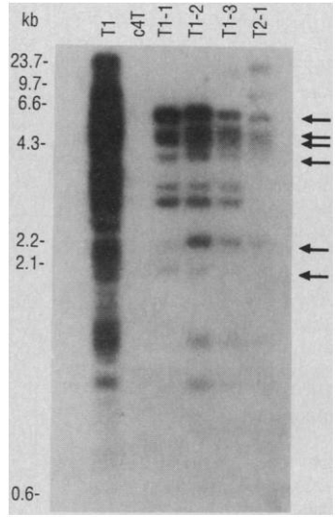

Fig. 1. DNA blot analysis of transfectants. Genomic DNA (10 μg) was digested with Pst I, subjected to agarose gel electrophoresis, and transferred to nitrocellulose. Hybridization was performed in 50% formamide, 5× SSC (standard saline citrate), 20 mM sodium phosphate (pH 6.5), 10% dextran sulfate, 1× Denhardt's solution, and denatured salmon sperm DNA (250 μg/ml) at 42°C. After hybridization, the filter was washed in 0.1× SSC (standard saline citrate) plus 0.1% SDS at 65°C. The probe was the Bam HI, Alu-containing fragment recloned from pBLUR8 into M13mp8 and isolated from the latter. Arrows indicate bands common to all secondary transfectants. Numbers at the left margin indicate sizes in kilobases.

Department of Pathology and Laboratory of Biomedical and Environmental Sciences, University of California, Los Angeles, CA 90024.

*The contribution of the first two authors should be considered equal.
†Present address: Department of Medical Oncology, City of Hope Medical Center, Duarte, CA 91090.
‡Present address: Exogene, Monrovia, CA 91016.
§To whom correspondence should be addressed.

characterization of corresponding cDNAs.

The c4T cell line, a hypoxanthine phosphoribosyl transferase–deficient derivative of the C^- mutant, c4, was cotransfected with plasmid pSV2gpt and genomic DNA from the human hepatoma cell line, HepG2, which expresses a functional Ah receptor. We inoculated treated cells into hypoxanthine-aminopterin-thymidine (HAT) medium—to select for cells expressing guanine phosphoribosyl transferase—and subjected the cells to the benzo[ghi]perylene plus near ultraviolet light "reverse selection" procedure (*11*) to select for cells that had reacquired AHH activity. Because the latter procedure provides only a transitory selection and because continuous HAT selection does not lead to stabilization of cotransfected genomic sequences in Hepa-1 cells (*12*), we subjected isolated clones at approximately biweekly intervals to the reverse selection in order to maintain selective pressure for the transfected *arnt*. Three clones, T1, T2, and T3 that survived repeated selection were obtained. Clones T2 and T3 could be sister clones; T1 was of independent origin. The cotransfection frequency [calculated as described in (*12*)] was 3×10^{-8}. Three secondary cotransfectants—T1-1, T1-2, and T1-3—were isolated after cotransfection of pSV2gpt and genomic DNA from the primary transfectant T1 into c4T cells and after exposure to HAT medium and the reverse selection procedure. In this instance, the cotransfection frequency was 2×10^{-8}. The secondary transfectant T2-1 was derived from the primary transfectant T2 by reverse selection only. Transfectants were capable of translocating the Ah receptor–dioxin complex to the nucleus and showed induction of AHH activity (Table 1).

DNA from transfectants was probed with the human Alu repetitive sequence BLUR8. All the primary transfectants contained large amounts of human DNA. The four secondary transfectants that were analyzed contained relatively small amounts of human DNA. Moreover, after digestion with Pst I, all secondary transfectants contained some BLUR8-hybridizing bands in common (Fig. 1); these bands therefore correspond to *arnt* or closely linked sequences.

Bgl II digestion of DNA from the secondary transfectant T1-1 generated BLUR8-hybridizing fragments of approximately 3, 5, 6, and 18 kb. A phage with an 18-kb Bgl II insert containing human DNA was isolated from a library prepared from T1-1 DNA that had been selected for fragment sizes between 16 and 20 kb. Analysis of this insert indicated that it did not contain the complete *arnt* sequence. Because we were also unable to demonstrate that this fragment contained any regions corresponding to transcribed sequences, we isolated contiguous genomic sequences by chromosome walking in a human genomic library. We found that a repeat-free fragment from one of these additional genomic clones hybridized to single restriction fragments in all mammalian species examined, suggesting that it contained a region corresponding to transcribed sequences. This fragment was then used to probe a cDNA library in λgt11 prepared from HepG2 mRNA. Two positive plaques were identified, and both contained the same 2.5-kb insert, which we termed M1-1.

DNA from HepG2 and c4T cells and the secondary transfectant T1-1 was probed with the cDNA clone M1-1. Hybridization and posthybridization were performed under low-stringency conditions. All hybridizing bands detected in HepG2 cells were also found in T1-1 cells. Therefore, except for the possibility of an extremely closely linked gene, no other gene appears to be closely related to *arnt* in the human genome. Additional experiments have shown that *arnt* is at least 60 kb in size.

Low-abundance *arnt* transcripts of 2.6 kb and 4.2 kb were detected in HepG2 cells, human liver, Hepa-1 cells, c4T, c39 (another C^- mutant), a B^- mutant, and a D^- mutant (Fig. 2). Although there were differences between the amounts of the *arnt* mRNA detected in Hepa-1 cells and the mutants in Fig. 2, such differences were not observed with additional RNA samples from these strains, and we conclude that no mutant has a significant abnormality in *arnt* mRNA expression. Both T1-1 and T2-1

Table 1. Characterization of transfectants. For the Ah receptor assay, cells were treated with 2 nM 2,3,7,8-tetrachloro-[^{3}H]dibenzo-*p*-dioxin (dioxin) (24 or 36 Ci/mmol) for 2 hours at 37°C. The Ah receptor was quantitated in both cytosolic and nuclear fractions after sucrose density gradient centrifugation (*9*). Cells for the AHH assay were grown in medium containing serum that had been treated with dextran-charcoal to remove inducers (*31*). AHH activity was assayed in untreated cells or in cells grown for 18 hours with 2 nM dioxin. The percentage of cells surviving reverse selection was ascertained by microscopic examination 4 hours after treatment with near ultraviolet light (*11*). The vT{1} cells were derived from c4T cells that had been treated with pBM5/NEO-M1-1; ~250 colonies that survived G418 treatment and the reverse selection were pooled, expanded in G418, and assayed. Cells of the vT{1} culture that survived a second application of the reverse selection procedure were pooled, expanded in G418, and assayed for the same parameters (vT{2}). The amounts of Ah receptor in cytosol and nucleus are in each case shown as a percentage of the value obtained for the same cellular compartment of Hepa-1 cells assayed at the same time. Where indicated, these values are the mean ± SD of three or four independent determinations, which in the case of vT{1} and of vT{2} were made at similar times in the culture history of the strains. The mean receptor levels in Hepa-1 cells (femtomoles per milligram of protein) are given in parentheses.

Strain	Ah receptor (% of Hepa-1)		AHH specific activity (% of Hepa-1 grown with dioxin)		Survival in reverse selection (%)
	Cytosol	Nucleus	− Dioxin	+ Dioxin	
Hepa-1	100 (140)	100 (267)	0.2	100	100
c4T	72 ± 35	2 ± 2	<0.2	<0.2	<0.0001
T1	89	56		21	
T1-1	94	90	0.2	36	
T2	96	53		30	
T2-1	105	45			
vT{1}	144 ± 22	18 ± 5	<0.1	9.2	15
vT{2}	71 ± 20	47 ± 17	0.2	34	40

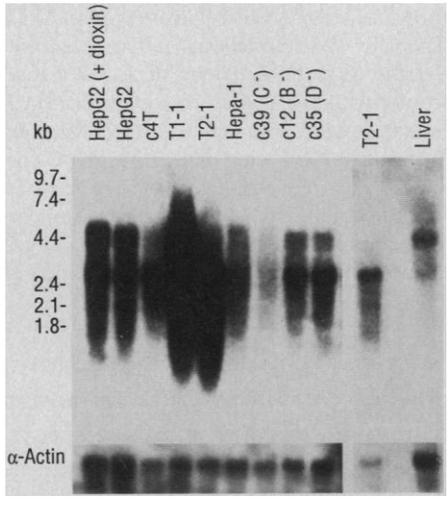

Fig. 2. RNA blot analysis of *arnt* mRNAs. Each lane contained 10 μg of polyadenylated RNA, except for the penultimate lane, which contained 0.5 μg. As indicated, one culture of HepG2 cells was grown for 24 hours with 2 nM dioxin. The filters were first probed with an Eco RI fragment from M1-1 corresponding to nucleotides 142 to 1203 of the cDNA sequence; they were then stripped and reprobed with a cDNA for mouse α-actin (lower part of figure). Polyadenylated RNA was prepared and analyzed by Northern blot as described (*31*), except that probes were ^{32}P-labeled to 10^9 cpm per microgram of DNA. The film was exposed at −70°C with intensifier screens for 18 hours for the M1-1 probe or 1 hour for the actin probe. Numbers at the left margin indicate sizes in kilobases.

secondary transfectants expressed considerably more of the 2.6-kb transcript than did the c4T mutant. T2-1 did not show increased expression of the 4.2-kb transcript, indicating that the shorter transcript is functional. The human Arnt protein may not be fully active in mouse cells and overexpression of *arnt* may have been necessary for transfectants of this gene to have survived the reverse selection. Neither of the two secondary transfectants shows amplification of either human or mouse *arnt*.

We used the M1-1 clone to screen two additional cDNA libraries prepared from HepG2 mRNA. Twenty-five clones were isolated and four were sequenced. The deduced amino acid sequence is shown in Fig. 3. The nucleotide sequences in the different clones were identical in regions of overlap, except for the 45-nucleotide (nt) segment and the differences in the sites of polyadenylation that are discussed below.

The cDNA clones analyzed represented both the longer and shorter transcripts. Clone M1-1 is derived from the latter, and has a tail of 14 adenines at its 3′ end. The cDNA sequence extends 2604 nt from the first base to the position corresponding to the start of the poly(A) tail in M1-1, indicating that little, if any, 5′ sequence is missing. An open reading frame extends from the start of the cDNA through position 2423. An ATG codon at position 57 is probably the translational initiation codon because it is bordered by nucleotides that conform to the consensus for translational initiation (*13*), and analysis of the sequence 5′ to this ATG with Staden programs for amino acid improbability (*14*) suggests that it is noncoding. M1-1 lacks a 45-nt segment that is present in two of the other cDNA clones, M2-1 and M3-1. This difference is probably due to alternative splicing because sequencing of *arnt* has shown that introns are located exactly at the 5′ and 3′ borders of the 45-nt segment (*15*).

The coding sequence of the cDNA was inserted into the mammalian expression vector pBM5/NEO. The presence of three internal Eco RI sites in the cDNA made it difficult to isolate the complete cDNA inserts from the larger clones, which had been isolated from λgt11 libraries. To obtain the desired cDNA segment, we subjected M1-1 to the polymerase chain reaction (PCR) in order to generate a product containing nt 46 to 2616 of the cDNA, but lacking the alternative exon. The PCR product was cloned into pBM5/NEO. Transfection of c4T cells with the resultant construct, pBM5/NEO-M1-1, generated G418-resistant clones at a frequency of 3×10^{-4} and clones able to survive reverse selection at a frequency of 3×10^{-3}. Transfection with the parental vector, pBM5/NEO, generated G418-resistant clones at a frequency of 6×10^{-4}, but clones able to survive the reverse selection at a frequency of less than 2×10^{-7} (*16*). The pBM5/NEO-M1-1–transfected cells that survived both selections were pooled and grown in G418-containing medium. The resulting culture, vT{1}, showed a partial restoration of nuclear translocation of the Ah receptor and possessed low AHH activity. These properties are explained by the observation that only 15% of the cells in the culture expressed AHH activity because only this percentage survived an additional exposure to the reverse selection procedure (Table 1).

The relatively poor phenotypic rescue by the cDNA expression plasmid is probably attributable to a number of factors, including disruption of the *arnt* cDNA during cellular processing of the transfected circular form, silencing of *arnt* cDNA expression, and loss of plasmids from the cells during the long period of growth required to generate enough cells for the Ah receptor assay (although at least one *neo* gene must have been retained in each cell). In this regard, high level expression of the human gene may be required to restore a substantial degree of receptor translocation to the mouse mutant; thus, many copies of the plasmid may be required per cell. The 15% of the vT{1} cells that survived the second exposure to the reverse selection procedure were expanded in G418. The degree of receptor translocation and the activity of AHH in the resulting culture, vT{2}, correlated with the fraction of cDNA-expressing cells (Table 1). The cDNA clone M3-1, which encodes an Arnt protein that lacks 163 COOH-terminal amino acids, was inserted into pBM5/NEO without the use of PCR (*17*). The resulting construct, pBM5/NEO-M3-1, behaved in a similar manner to pBM5/NEO-M1-1 in the above assays, indicating that sequence errors

```
MAATTANPEM TSDVPSLGPA IASGNSGPGI QGGGAIVQRA IKRRPGLDFD  50
DDGEGNSKFL RCDDDQMSND KERFARSDDE QSSADKERLA RENHSEIERR 100
RRNKMTAYIT ELSDMVPTCS ALARKPDKLT ILRMAVSHMK SLRGTGNTST 150
DGSYKPSFLT DQELKHLILE AADGFLFIVS CETGRVVYVS DSVTPVLNQP 200
QSEWFGSTLY DQVHPDDVDK LREQLSTSEN ALTGRILDLK TGTVKKEGQQ 250
SSMRMCMGSR RSFICRMRCG SSSVDPVSVN RLSFVRNRCR NGLGSVKDGE 300
PHFVVVHCTG YIKAWPPAGV SLPDDDPEAG QGSKFCLVAI GRLQVTSSPN 350
CTDMSNVCQP TEFISRHNIE GIFTFVDHRC VATVGYQPQE LLGKNIVEFC 400
HPEDQQLLRD SFQQVVKLKG QVLSVMFRFR SKNQEWLWMR TSSFTFQNPY 450
SDEIEYIICT NTNVKNSSQE PRPTLSNTIQ RPQLGPTANL PLEMGSGQLA 500
PRQQQQQTEL DMVPGRDGLA SYNHSQVVQP VTTTGPEHSK PLEKSDGLFA 550
QDRDPRFSEI YHNINADQSK GISSSTVPAT QQLFSQGNTF PPTPRPAENF 600
RNSGLAPPVT IVQPSASAGQ MLAQISRHSN PTQGATPTWT PTTRSGFSAQ 650
QVATQATAKT RTSQFGVGSF QTPSSFSSMS LPGAPTASPG AAAYPSLTNR 700
GSNFAPETGQ TAGQFQTRTA EGVGVWPQWQ GQQPHHRSSS SEQHVQQPPA 750
QQPGQPEVFQ EMLSMLGDQS NSYNNEEFPD LTMFPPFSE             789
```

Fig. 3. Predicted sequence of the Arnt protein. The alternative exon is underlined. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. The nucleotide sequence of *arnt* has been submitted to GenBank.

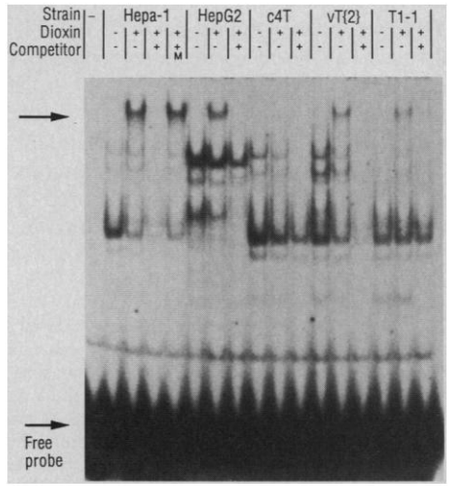

Fig. 4. Gel retardation analysis of transfectants. Cells were grown with serum that had been treated with dextran-charcoal and were incubated with or without 2 nM dioxin for 1 hour before being harvested for analysis. The gel retardation assay was performed as described (*32*), except that the double-stranded, synthetic oligonucleotide probe had the sequence 5′-CCTCCAGGCTCT-TCTCACGCAACTCCGGGGCAC-3′ (coding strand), which corresponds to nucleotides −1005 to −973, relative to the cap site of the mouse P450IA1 gene. The probe was labeled with [λ-^{32}P]adenosine triphosphate with the use of polynucleotide kinase to a specific activity of 1.45×10^6 cpm/pmol. Incubations contained 5 μg of nuclear protein. As indicated, some incubations also contained a 200-fold excess of the unlabeled oligonucleotide (competitor) or a 200-fold excess of an unlabeled mutant oligonucleotide (M), which had the same sequence except that 5′-TA-3′ replaced 5′-CG-3′ at positions −988 and −987. The positions of the free probe and receptor-dependent retarded band are indicated.

generated by PCR cannot explain the submaximal results obtained with pBM5/NEO-M1-1 and that both the alternative exon and the 163 COOH-terminal amino acids of the Arnt protein are dispensable.

Gel retardation experiments with nuclear extracts prepared from Hepa-1 cells grown in dioxin have identified a protein-XRE complex. The retarded complex has been shown to contain the ligand-binding subunit of the Ah receptor (*2*). The C^- mutant c4T does not generate this complex. We performed a gel retardation experiment with the expression vector–derived transfectant culture vT{2} and with the genomic DNA–derived secondary transfectant T1-1 to see whether formation of the XRE–Ah receptor complex was restored in the transfectants. A 33-bp double-stranded oligonucleotide that encompasses XRE 1 was used as the probe. A dioxin-inducible retarded band was seen with Hepa-1 and HepG2 cells (Fig. 4); this band was eliminated by competition with a 200-fold excess of unlabeled oligonucleotide but not by competition with an oligonucleotide containing two nucleotide substitutions in the core sequence of XRE 1. The band was not seen with c4T extracts, thus providing evidence that it corresponds to the XRE–Ah receptor complex. The complex was detected in extracts of vT{2} and T1-1 cells, but only if the cells had been grown with dioxin. The intensity of the retarded band was less in both types of transfectant than with either Hepa-1 or HepG2 cells. However, when they were harvested for this assay, vT{2} and T1-1 cells expressed only 75% and 36%, respectively, of the dioxin-induced AHH activity that was present in Hepa-1 cells. These experiments confirm that the *arnt* cDNA restores Ah receptor activity—both with regard to its nuclear translocation and with regard to its subsequent binding to the XRE—to the C^- mutant.

The vector pBM5/NEO contains an SV40 origin of replication and should therefore replicate in monkey COS-7 cells. COS-7 cells transiently transfected with pBM5/NEO-M1-1 and pBM5/NEO-M3-1 contained much larger concentrations of *arnt* mRNA than cells transfected with the parental plasmid. However, when analyzed for ^{3}H-labeled dioxin binding by an assay optimized for the human Ah receptor (*18*), cells transfected with the M1-1 and M3-1 cDNA constructs possessed Ah receptor levels of 820 and 773 fmol per milligram of protein, respectively, which were no greater than the level in cells transfected with the parental plasmid (774 fmol/mg), indicating that the Arnt protein does not bind dioxin.

The predicted size of the Arnt protein is 86,637 daltons (789 amino acids) including, and 85,035 daltons (774 amino acids) excluding, the alternative exon. No hydrophobic leader sequence or potential transmembrane domains are apparent. There is a cysteine-rich area between amino acids 256 and 400, which has no similarity to the Zn^{2+}-finger domains of steroid hormone receptors (*19*).

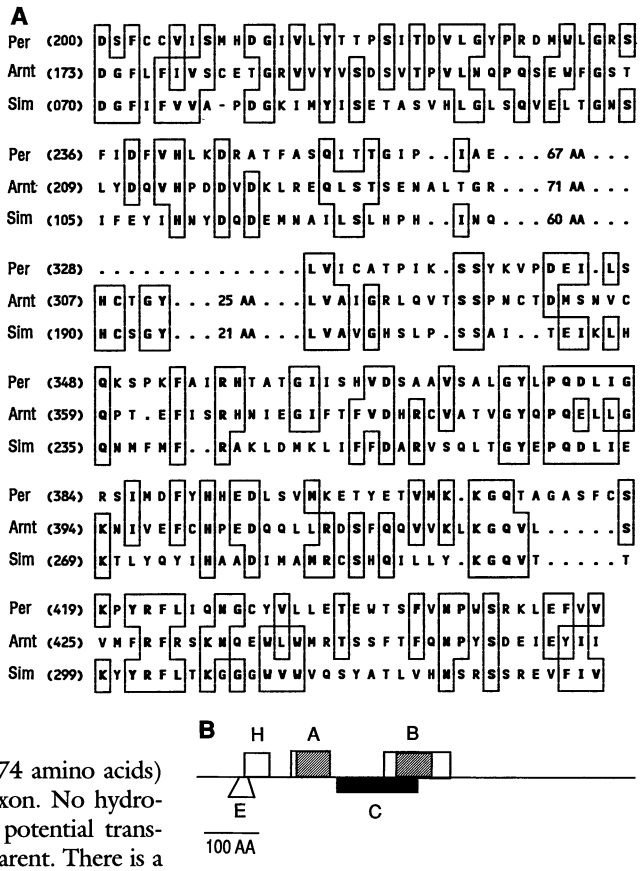

Fig. 5. (**A**) Alignment of residues 173 to 458 of the Arnt protein with residues 200 to 452 of Per and residues 70 to 332 of Sim. Identities are boxed. (**B**) Schematic of the Arnt protein. Boxes indicate the locations of the similarities to the basic helix-loop-helix motif consensus sequence (H) and the Per and Sim proteins (A and B). The ~50–amino acid repeats are hatched. The position of the alternative exon (E) is indicated in box H. The solid bar (C) beneath the figure represents the cysteine-rich region. AA, amino acids.

A database search revealed no sequences having similarity to the *arnt* cDNA sequence. The Arnt amino acid sequence has two regions that are similar to both the *Drosophila* circadian rhythm protein (Per) and the *Drosophila* single-minded protein (Sim) (Fig. 5A). Profile analysis (*20*) of the alignment of the first segment (amino acids 173 to 235) of Arnt with Per and Sim gave scores that were 4.45 and 4.25 SD, respectively, above the average score of the alignment to the protein databases. The second segment (amino acids 337 to 458) of similarity had alignment scores of 9.7 SD (Arnt-Per) and 5.6 SD (Arnt-Sim) above the average score. Each area of similarity in Arnt contains one copy of the ~50–amino acid repeat previously described in the *Drosophila* proteins (*21*). No function has been assigned to these two segments in the *Drosophila* proteins.

Another segment of Arnt shows high conformity to the consensus sequence for the basic helix-loop-helix motif of certain DNA binding proteins (*22*). This segment of similarity in Arnt encompasses the alternative exon, but the similarity is apparent regardless of whether or not this exon is included. The segment in Arnt comprises amino acids 74 to 145, if the alternative exon is excluded, or amino acids 89 to 145, if the alternative exon is included. The domain structure of the Arnt protein is shown in Fig. 5B.

In conclusion, it is likely that *arnt* corresponds to the gene that is mutated in the C^- mutants rather than a suppressor gene because it is capable of repairing all aspects of the C^- mutant phenotype that we have assayed, and, furthermore, it is expressed in both C^- mutant cells and Hepa-1 cells. However, further experments will be required to provide definitive proof of this. The mouse *arnt* gene is located on chromosome 3 (*23*), whereas the gene for the ligand-binding subunit of the Ah receptor is located on mouse chromosome 12 (*24*). These findings are consistent with our inability to detect [^{3}H]dioxin binding by the *arnt* cDNA–encoded protein, indicating that Arnt is not the ligand-binding subunit of the Ah receptor. Arnt has no sequence similarity with Hsp90, which is associated with the Ah receptor, or with Hsp70, which has been found associated with the progesterone receptor (*25*). Whether Arnt truly directs the Ah receptor to translocate from the cytosol to the nucleus after binding of ligand or whether the unoccupied receptor is, in fact, located in the nucleus and Arnt increases the avidity with which the receptor binds to this organelle requires further investigation. In any event, the mode of action of the Ah

receptor appears to be more complex than that of the steroid hormone receptors, and particularly the glucocorticoid receptor: this is indicated by the observations that (i) the c4T mutant is not defective in ligand-dependent nuclear translocation of the glucocorticoid receptor, which suggests that Arnt is not required for this process (*26*); and (ii) somatic cell mutants defective in ligand binding or nuclear translocation of the glucocorticoid receptor fall within the same complementation group, which corresponds to the glucocorticoid receptor structural gene (*27*), whereas analogous mutants for the Ah receptor fall into three complementation groups.

The sequence similarity between the Arnt protein and the *Drosophila* Per and Sim proteins is of interest because Sim is a nuclear protein (*21*), whereas Per is found in either the nucleus or cytoplasm, depending on the tissue examined (*28*). The basic helix-loop-helix motif of Arnt is found in certain proteins that bind DNA as homodimers or heterodimers. The motif contains domains for DNA binding and for dimerization (*29*).

Humans are heterogeneous with regard to the degree of inducibility or the maximally induced activity of cytochrome P450IA1–dependent AHH in their lymphocytes and monocytes. This heterogeneity appears to have a genetic origin, and, furthermore, individuals with high AHH activity appear to be more susceptible to cigarette-induced lung cancer (*30*). Whether differences in *arnt* structure, expression, or both underlie this heterogeneity in the human population will be of interest.

REFERENCES AND NOTES

1. A. Fujisawa-Sehara, K. Sogawa, M. Yamane, Y. Fujii-Kuriyama, *Nucleic Acids Res.* **15**, 4179 (1987); M. S. Denison, J. M. Fisher, J. P. Whitlock, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2528 (1988).
2. M. S. Denison, J. M. Fisher, J. P. Whitlock, Jr., *J. Biol. Chem.* **263**, 17221 (1988); A. Fujisawa-Sehara, M. Yamane, Y. Fujii-Kuriyama, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5859 (1988); J. Hapgood, S. Cuthill, M. Denis, L. Poellinger, J.-A. Gustafsson, *ibid.* **86**, 60 (1989).
3. D. M. Jerina, *Drug Metab. Dispos.* **11**, 1 (1983); D. W. Nebert and F. J. Gonzalez, *Annu. Rev. Biochem.* **56**, 945 (1987); H. Ohgaki *et al., Environ. Health Perspect.* **67**, 129 (1986); M. E. McManus *et al., Carcinogenesis* **10**, 357 (1989); M. Degawa, S. Tanimura, T. Agatsuma, Y. Hashimoto, *ibid.*, p. 1119; D. Rodrigues *et al., Eur. J. Biochem.* **181**, 627 (1989); E. G. Snyderwine and N. Battula, *J. Natl. Cancer Inst.* **81**, 223 (1989); H. Wallin *et al., Carcinogenesis* **11**, 489 (1990).
4. A. Poland, in *Biological Mechanisms of Dioxin Action,* A. Poland and R. D. Kimbrough, Eds. (Banbury Report No. 18, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1984), pp. 109–117; S. H. Safe, *Annu. Rev. Pharmacol. Toxicol.* **26**, 371 (1986).
5. M. S. Denison, L. M. Vella, A. B. Okey, *J. Biol. Chem.* **261**, 3987 (1986); S. Cuthill, L. Poellinger, J.-A. Gustafsson, *ibid.* **262**, 3477 (1987); G. H. Perdew and A. Poland, *ibid.* **263**, 9848 (1988); G. H. Perdew, *ibid.*, p. 13802; M. Denis, S. Cuthill, A.-C. Wikström, L. Poellinger, J.-A. Gustafsson, *Biochem. Biophys. Res. Commun.* **155**, 801 (1988).
6. J. P. Whitlock, Jr., and D. R. Galeazzi, *J. Biol. Chem.* **259**, 980 (1984); M. S. Denison, P. A. Harper, A. B. Okey, *Eur. J. Biochem.* **155**, 223 (1986); J. M. Gudas, S. O. Karenlampi, O. Hankinson, *J. Cell Physiol.* **128**, 441 (1986); A. Wilhelmsson *et al., EMBO. J.* **9**, 69 (1990).
7. J.-A. Gustafsson *et al., Endocrine Rev.* **8**, 185 (1987); E. C. Henry, G. Rucci, T. A. Gasiewicz, *Biochemistry* **28**, 6430 (1989).
8. O. Hankinson, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 373 (1979); in *Microsomes, Drug Oxidations, and Chemical Carcinogenesis,* M. J. Coon *et al.*, Eds. (Academic Press, New York, 1980), pp. 1149–1152; *Somat. Cell Genet.* **7**, 373 (1981); *ibid.* **9**, 497 (1983).
9. S. O. Karenlampi, C. Legraverend, J. M. Gudas, N. Carramanzana, O. Hankinson, *J. Biol. Chem.* **263**, 10111 (1988).
10. C. Legraverend *et al., ibid.* **257**, 6402 (1982).
11. J. R. Van Gurp and O. Hankinson, *Cancer Res.* **43**, 6031 (1983).
12. D. F. Montisano and O. Hankinson, *Mol. Cell Biol.* **5**, 698 (1985).
13. M. Kozac, *Cell* **44**, 283 (1986); J. Kozak, *Nucleic Acids Res.* **12**, 857 (1984).
14. D. McLachlan, R. Staden, D. R. Boswell, *Nucleic Acids Res.* **12**, 9567 (1984).
15. E. C. Hoffman and O. Hankinson, unpublished data.
16. PCR was performed on M1-1, which extends from nt 74 to 2616, with the use of an upstream primer with the sequence 5′-CCCGGATCC-3′ followed by nt 46 to 117 of the cDNA sequence. The downstream primer contained the sequence 5′-CCCGGATCC-3′ followed by the 24 nt of λgt11 immediately adjacent to the Eco RI cloning site. Both primers therefore contain Bam HI sites. The PCR product was digested with Bam HI and ligated into the Bam HI site of pcDNA-1 (Invitrogen, San Diego, CA) to generate pcDNA-1-M1-1. The insert was excised from pcDNA-1-M1-1 with Sac I and Not I and was ligated to Sac I– and Not I–digested pBM5/NEO (Invitrogen) in order to generate a clone containing the insert in the correct orientation for expression. Colonies containing the insert were analyzed by restriction mapping for the presence of the 95-bp Sac I–Sac I fragment that lies between the cytomegalovirus promoter and the M1-1 insert. The c4T cells were transfected with 20 μg of pBM5/NEO or pBM5/NEO-M1-1. Three days later, cells were selected in G418 (400 μg/ml), subjected to reverse selection, or subjected to both selections.
17. M3-1 extends from position 1 to position 1935 and was obtained from a λ ZAP library (Stratagene). The cDNA was excised from this vector by a Bam HI and Hind III double digestion and ligated into pBM5/NEO.
18. D. K. Manchester, S. K. Gordon, C. L. Golas, E. A. Roberts, A. B. Okey, *Cancer Res.* **47**, 4861 (1987).
19. R. M. Evans, *Science* **240**, 889 (1988).
20. M. Gribskov, M. Homyak, J. Edenfield, D. Eisenberg, *CABIOS* **4**, 61 (1988).
21. S. T. Crews, J. B. Thomas, C. S. Goodman, *Cell* **52**, 143 (1988).
22. R. Benezra, R. L. Davis, D. Lockshon, D. L. Turner, H. Weintraub, *ibid.* **61**, 49 (1990).
23. B. Brooks *et al., Am. J. Hum. Genet.* **45** (suppl. 4), A132 (1989).
24. A. Poland, E. Glover, B. A. Taylor, *Mol. Pharamacol.* **32**, 471 (1987).
25. S. L. Kost, D. F. Smith, W. P. Sullivan, W. J. Welch, D. O. Toft, *Mol. Cell. Biol.* **9**, 3829 (1989).
26. O. Hankinson and R. M. Bannister, in *Microsomes and Drug Oxidations,* J. Miners, D. J. Birkett, R. Drew, B. K. May, M. McManus, Eds. (Taylor and Francis, London, 1988), pp. 39–46.
27. U. Gehring, K. R. Yamamoto, G. M. Tomkins, *Res. Steroids* **7**, 43 (1977); M. Pfahl and S. Bourgeois, *Somat. Cell Genet.* **6**, 63 (1980); R. Miesfeld *et al., Nature* **312**, 779 (1984); M. Danielsen, J. P. Northrop, G. M. Ringold, *EMBO J.* **5**, 2513 (1986).
28. L. Saez and M. W. Young, *Mol. Cell. Biol.* **8**, 5378 (1988).
29. H. Weintraub *et al., Science* **251**, 761 (1991).
30. R. E. Kouri *et al., Cancer Res.* **42**, 5030 (1982); T. Okuda *et al., ibid.* **37**, 3904 (1977); H. V. Gelboin, *N. Engl. J. Med.* **309**, 105 (1983).
31. O. Hankinson *et al., J. Biol. Chem.* **260**, 1790 (1985).
32. F. Saatcioglu, D. J. Perry, D. S. Pasco, J. B. Fagan, *ibid.* **265**, 9251 (1990).

33. We thank B. Lovejoy, S. Eisenberg, S. Crews, and S. Kohtz for help with the sequence analysis; A. Poland and S. Safe for [^{3}H]dioxin; R. H. Tukey for human liver mRNA; A. J. Lusis and F. S. Hagen for cDNA libraries; A. J. Watson for advice on the gel retardation experiment; J. B. Fagan for information before publication; A. B. Okey and R. D. Prokipcak for advice; and M. Pastor for typing the manuscript. Supported by NCI grants CA 28868 and CA 16048, Department of Energy contract DE FC03 87ER 60615, a predoctoral stipend (E.C.H.) from Associated Western Universities, a predoctoral stipend (H.R.) from American Cancer Society Institutional Grant IRG-131, and a postdoctoral stipend (B.A.B.) from the California Institute for Cancer Research.

10 July 1990; accepted 15 February 1991

Site-Specific Recombination Between Homologous Chromosomes in *Drosophila*

KENT G. GOLIC*

The ability to mark a cell and its descendants genetically so that the resulting cell clone can be distinguished from neighboring cells facilitates studies in animal biology and development. A method of generating clones by inducing homologous mitotic recombination in *Drosophila* with a site-specific yeast recombinase is described. This method allows for frequent mosaicism after mitotic exchange is induced at predefined sites in the genome.

MOSAIC ORGANISMS—THOSE WITH cells of different genotypes in a single individual—have provided insights into such issues as the developmental potential and determination of cell clones, cell lineage relations, and restrictions in development, patterns of cell division, and cell autonomy (*1*). In *Drosophila*, such studies have been greatly aided by the availability of several methods for producing these genetic mosaics. Early in development, mosaics can be produced with the use of mutations that cause chromosome instability or with special ring chromosomes that are inherently unstable (*2*). Later in devel-