Characterization of Multiple Forms of the Ah Receptor: Recognition of a Dioxin-Responsive Enhancer Involves Heteromer Formation†

Thomas A. Gasiewicz,*,‡ Cornelis J. Elferink,§ and Ellen C. Henry‡

Department of Biophysics, Environmental Health Sciences Center, University of Rochester School of Medicine, Rochester, New York 14642, and Department of Pharmacology, Stanford University School of Medicine, Stanford, California 94305

Received September 12, 1990; Revised Manuscript Received November 19, 1990

ABSTRACT: We have employed a combination of gel retardation, protein–DNA cross-linking, and protein–protein cross-linking techniques to further examine the 2,3,7,8-tetrachlorodibenzo-*p*-dioxin- (TCDD-) dependent changes in the Ah receptor that result in a DNA-binding conformation. Gel retardation analysis of DNA–Sepharose chromatographic fractions of rat hepatic cytosol indicated that TCDD-dependent and sequence-specific DNA binding coeluted with a 200-kDa form of the Ah receptor (peak 2) previously characterized as being multimeric and having high affinity for calf thymus DNA. The TCDD-bound, 100-kDa form of the receptor (peak 1) bound weakly to the DNA recognition motif. These results indicated that the DNA-binding form of the Ah receptor is a multimer. SDS–polyacrylamide gel electrophoresis of peak 2 cross-linked to a bromodeoxyuridine-substituted DNA recognition motif indicated that this form of the receptor present in rat hepatic cytosol is composed of at least two DNA-binding proteins of approximately 100 and 110 kDa. Using the chemical cross-linking agent dimethyl pimelimidate, we further established that the 100-kDa form of the receptor (peak 1) associates with a different protein to generate the receptor form (peak 2) that binds to the dioxin-responsive enhancer. Photoaffinity-labeling studies indicated that only the 100-kDa protein (peak 1), and not the 110-kDa protein, binds ligand. Together, these observations imply that the DNA-binding form of the Ah receptor exists as a heteromer.

The Ah (*a*ryl *h*ydrocarbon) receptor is a soluble intracellular protein that binds TCDD[1] and a number of related xenobiotics with specificity and high affinity. The ability of the ligand–Ah receptor complex to associate with specific nuclear sites and modulate gene transcription has been implicated in the mechanism whereby these compounds elicit a variety of tissue-specific biochemical and toxic responses in exposed animals (Poland & Knutson, 1982; Whitlock, 1990). Research to date has suggested that the Ah receptor is both physicochemically and functionally analogous in many respects to the steroid hormone receptors (Wilhelmsson et al., 1986; Cuthill et al., 1987; Gustafsson et al., 1987). However, no endogenous ligand for the Ah receptor has yet been identified, and its exact role in the regulation of cellular processes is unknown.

Like the steroid hormone receptors, the Ah receptor undergoes a ligand-dependent transformation to a form having high affinity for DNA (Hannah et al., 1986; Gasiewicz & Bauman, 1987). We have previously described a model for Ah receptor transformation in which exposure of the DNA-binding site on the receptor, under conditions in vitro, initially involves a ligand-dependent dissociation of some component or components from the unoccupied form, having an approximate molecular mass of 300 kDa, to yield a smaller species of approximately 100 kDa (Henry et al., 1989). The 90-kDa heat shock protein is a likely candidate for a dissociable component since it is associated with a large molecular mass form of the Ah receptor (Perdew, 1988; Denis et al., 1988) and inhibits its DNA binding (Wilhelmsson et al., 1990). Although the 100-kDa form appears to be the monomeric ligand-binding subunit of the Ah receptor (Poland & Glover, 1987, 1988; Henry et al., 1989), an additional form with a molecular mass of 160–200 kDa has higher affinity for calf thymus DNA and exhibits hydrodynamic similarities to the receptor isolated from nuclei of animals treated with TCDD (Henry et al., 1989; Prokipcak & Okey, 1988). It is not known whether this form is simply a dimer of the monomeric subunit or a result of the association of this subunit with another protein or proteins.

We have previously characterized different forms of the Ah receptor partially on the basis of their ability to bind nonspecifically to calf thymus DNA. Although the interaction of trans-acting factors with nonspecific DNA in vivo is important in both facilitating and buffering binding to specific DNA sequences (Winter & von Hipple, 1981), sequence-specific binding is ultimately crucial to transcriptional regulation by the receptor. Specific genomic regulatory elements (dioxin-responsive enhancers) that mediate the transcriptional response to the transformed Ah receptor have been identified upstream of the structural gene for CYPIA1 (Jones et al., 1985; Gonzalez & Nebert, 1985; Fujisawa-Sehara et al., 1986). TCDD–receptor complexes in crude cytosol (Fujisawa-Sehara et al., 1988; Nemoto et al., 1990) and nuclear extracts (Denison et al., 1989; Shen & Whitlock, 1989) recognize the core sequence $^{5'}$T-GCGTG-$^{3'}$ / $^{3'}$A-CGCAC-$^{5'}$ that is present in four copies within the enhancer element (Denison et al., 1989; Fisher et al., 1990). Knowledge of this receptor–enhancer interaction is necessary for understanding both the ligand-dependent molecular changes in the receptor that result in a

†Supported by NIH Grant ES02515, Center Grant ES01247, and a grant from the International Life Sciences Research Institute Foundation to T.A.G. C.J.E. was supported in part by NIH Grant ES03719.

*Address correspondence to this author.

‡University of Rochester School of Medicine.

§Stanford University School of Medicine.

[1] Abbreviations: TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; TCDF, 2,3,7,8-tetrachlorodibenzofuran; IBr_2DD, 2-iodo-7,8-dibromodibenzo-*p*-dioxin; BSA, bovine serum albumin; DMP, dimethyl pimelimidate; CYPIA1, cytochrome P450IA1.

0006-2960/91/0430-2909$02.50/0

DNA-binding conformation and the role of the Ah receptor in transcriptional regulation. We report here that the form of the Ah receptor contained in cytosol that has the highest affinity for calf thymus DNA also binds to the recognition motif contained within the enhancer element, and does so as a heteromer.

Experimental Procedures

Chemicals. [^{3}H]TCDD (30 Ci/mmol) was the gift of Dr. S. Safe (Texas A&M University) and [^{125}I]-2-iodo-7,8-dibromodibenzo-*p*-dioxin ([^{125}I]IBr$_2$DD) (2176 Ci/mmol) was a gift from Dr. A. Poland (University of Wisconsin). [α-^{32}P]dATP (3000 Ci/mmol) was from Amersham Corp. 5-Bromo-2′-deoxyuridine-5′-triphosphate (BrdU) was from Pharmacia LKB. Calf thymus DNA and cyanogen bromide activated Sepharose 4B were from Sigma. Molecular biological reagents were from Bethesda Research Laboratories, Boehringer Mannheim, and New England Biolabs.

Buffers. HEDG buffer consisted of 25 mM Hepes, 1.5 mM EDTA, 1 mM dithiothreitol, and 10% (v/v) glycerol, pH adjusted to 7.6 at room temperature. For some procedures, this buffer was modified as noted in the figures.

Animals and Tissue Preparation. Male Sprague-Dawley rats (200–250 g) were from either Simonsen Laboratories (Gilroy, CA) or Charles River (Wilmington, MA). Syrian golden hamsters (110–150 g) and Hartley guinea pigs (300–350 g) were from Charles River. Animals were housed with a 12-h light–dark cycle and were allowed food and water ad libitum. Hepatic cytosol was prepared as described previously (Gasiewicz & Bauman, 1987) and adjusted to a protein concentration of approximately 15 mg/mL. [^{3}H]TCDD or [^{125}I]IBr$_2$DD (in dimethyl sulfoxide or *p*-dioxane) was added to cytosol at a final concentration of 2 nM, and the cytosol was incubated for 2 h at 20 °C or 18 h at 0–4 °C. Both incubation conditions produce maximal transformation of the Ah receptor in vitro (Gasiewicz & Bauman, 1987; Henry et al., 1989). Aliqouts of transformed cytosol were stored at −80 °C and were analyzed within 3 weeks.

DNA–Sepharose Chromatography and Partial Purification of the Transformed Ah Receptor. Preparation of DNA–Sepharose, column packing, chromatographic conditions, elution with a NaCl gradient, and analysis of the eluted fractions were as described by Gasiewicz and Bauman (1987), with the exceptions that the bed volume of the column was approximately 20 mL insteady of 10 mL and the column was equilibrated with HEDG buffer containing 0.05 M NaCl rather than 0.01 M NaCl. The high-affinity DNA-binding form of the receptor (160–200 kDa) was partially purified by a combination of ammonium sulfate precipitation and DNA–Sepharose chromatography. Crude hepatic cytosol was incubated with [^{3}H]TCDD as described above. A saturated ammonium sulfate solution (in 25 mM Hepes, pH 7.6, 1.5 mM EDTA) was added dropwise to the cytosol with constant stirring at 4 °C to a final ammonium sulfate concentration of 33.3%. The resulting suspension was incubated for 15 min at 4 °C prior to centrifugation at 9000 rpm (11000g_{av}) for 30 min at 4 °C. The supernate was discarded and the pellet was washed once with 10 mL of 33.3% saturated ammonium sulfate solution. Following a second centrifugation and removal of the supernate, the pellet was dissolved in a volume of HEDG buffer equivalent to about 15% of the original cytosol volume. The ammonium sulfate precipitation removed approximately 90% of the cytosolic protein while concentrating approximately 80% of the specific binding of [^{3}H]TCDD as determined by the hydroxylapatite binding assay (Gasiewicz & Neal, 1982) (data not shown). The redissolved pellet was loaded onto a DNA–Sepharose column, and specific binding was eluted as described previously (Henry et al., 1989). Eluted fractions that contained the high-affinity DNA-binding form of the receptor (peak 2, see Figure 1) were pooled, concentrated by the ammonium sulfate procedure described above, redissolved in HEDG buffer containing BSA (4 mg/mL), and frozen at −80 °C. The combination of ammonium sulfate precipitation and DNA–Sepharose chromatography yielded a preparation containing about 40% of the original amount of [^{3}H]TCDD specific binding with a purification of approximately 1500-fold.

The ligand-bound 100-kDa (peak 1) and 160–200-kDa (peak 2) forms of the receptor, which have low and high affinities for calf thymus DNA, respectively, were partially purified (approximately 250-fold) from guinea pig, hamster, or rat hepatic cytosol by DNA–Sepharose chromatography as previously described (Henry et al., 1989). Aliquots of these were frozen at −80 °C in the presence of BSA (4 mg/mL).

Synthetic Oligodeoxyribonucleotides. The complementary oligodeoxyribonucleotides 5′-GATCCGGCTCTTCTCACGCAACTCCGAGCTCA-3′ and 5′-GATCTGAGCTCGGAGTTGCGTGAGAAGAGCCG-3′, which when annealed contain a single core recognition sequence (underlined) for the DNA-binding form of the Ah receptor, were synthesized and ^{32}P-labeled at the 5′ ends as previously described (Denison et al., 1988). For the protein–DNA cross-linking studies, thymidine residues were replaced by bromodeoxyuridines (Lin & Riggs, 1974). The first oligonucleotide described above was annealed to the 8-nucleotide primer 5′-TGAGCTCG-3′, and 1 μg of the primed template was end-filled by the Klenow fragment of DNA polymerase at 37 °C for 1 h in a reaction volume of 20 μL, containing 50 μM each dCTP, dGTP, 5-bromo-2′-deoxyuridine triphosphate and 5 μM [α-^{32}P]dATP (Maniatis et al., 1982). Non-radiolabeled BrdU-substituted DNA was synthesized by substituting 50 μM dATP for the [α-^{32}P]ATP. Analogous complementary synthetic oligodeoxyribonucleotides containing 5-methylcytosine (m^{5}C) in the sequences 5′-GATCCGGCTCTTGTCAm^{5}CGCAACTCCGAGCTCA-3′ and 5′-GATCTGAGCTCGGAGTTGm^{5}CGTGAGAAGAGCCG-3′ were prepared as described previously (Shen & Whitlock, 1989).

Gel Retardation Assay. Cytosolic protein (60–100 μg) or an aliquot (3–4 μL) of the chromatography fractions or the partially purified receptor preparation was mixed with 250 ng, 25 ng, or 100 ng of poly[d(I-C)], respectively, and 5×10^4–10^5 cpm of the DNA probe (specific activity 10^8 cpm/μg of DNA) and was analyzed by nondenaturing gel electrophoresis as previously described (Denison et al., 1988).

Protein–Protein Cross-Linking. Liganded receptor peaks 1 and 2 isolated by DNA–Sepharose chromatography were treated with the cross-linking agent dimethyl pimelimidate (DMP) (20 mM) as described by Aranyi et al. (1988). Following cross-linking at 10 °C for 30 min, [^{3}H]TCDD or [^{125}I]IBr$_2$DD was covalently bound to the receptor by UV irradiation (maximun wavelength 300 nm, 10 min, 0 °C) (Landers et al., 1989; Poland et al., 1986). The protein was precipitated with ice-cold acetone overnight at −20 °C, dissolved in loading buffer [62.5 mM Tris-HCl, pH 7.0, 2.3% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 10% (v/v) glycerol] and electrophoresed on a SDS–polyacrylamide gel (4.5% stacking gel, 6% separating gel) (Laemmli, 1970). The ^{125}I-labeled proteins were visualized by autoradiography. Gels containing the ^{3}H-labeled proteins were sliced into 5-mm sections and dissolved in 1 mL of NCS tissue solubilizer (Amersham Corp.). Radioactivity was determined by using

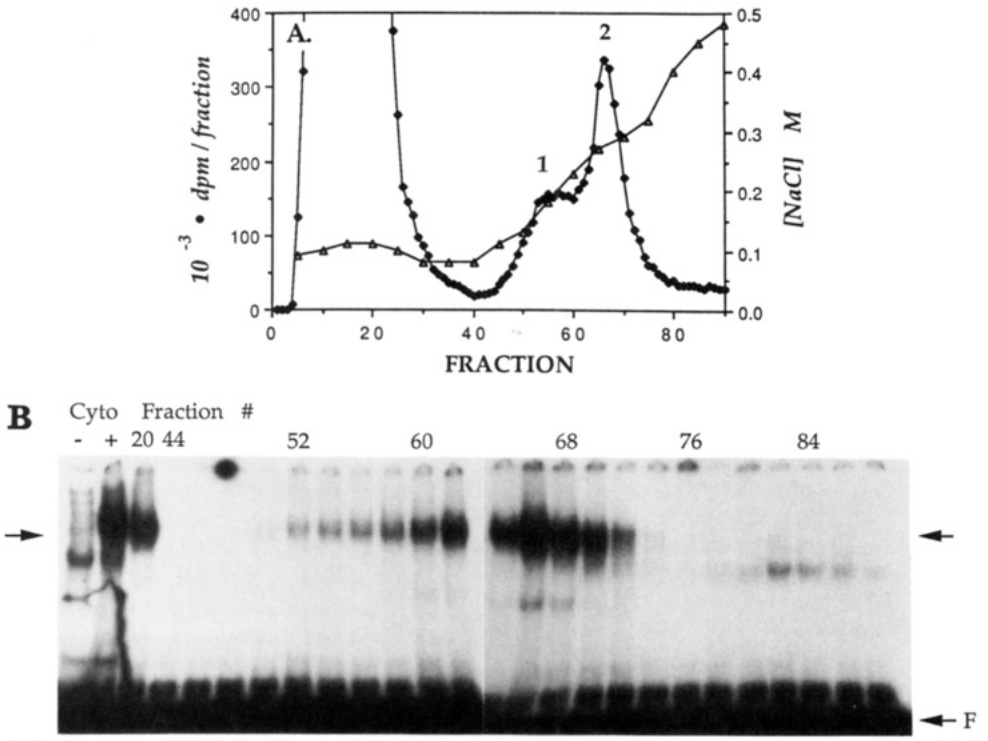

FIGURE 1: Gel retardation analysis of DNA–Sepharose chromatography fractions. Rat hepatic cytosol (~15 mg of protein/mL) was incubated with [^{3}H]TCDD for 2 h at 20 °C. The 33% ammonium sulfate pellet was prepared, dissolved in HEDG buffer, and chromatographed on DNA–Sepharose. Aliquots of eluted fractions were used for counting radioactivity and determining conductivity (A) and for gel retardation assay (B). Numbers above lanes in (B) indicate column fraction numbers; Cyto = crude cytosol incubated in the absence (−) or presence (+) of TCDD; F = free ^{32}P-labeled oligonucleotide; the upper arrow indicates TCDD-dependent bands.

a Packard Tri-Carb Model 4530 scintillation spectrometer. The counting efficiency for tritium was 40–60%.

Protein–DNA Cross-Linking. The gel retardation mixture containing the partially purified DNA-binding form of the receptor and the ^{32}P-labeled BrdU-containing DNA was irradiated at 302 nm (maximum emission wavelength; maximum intensity, 7000 μW/cm^2) for 15 min at 0 °C and electrophoresed on either a 4% polyacrylamide gel (Denison et al., 1988) or a SDS–polyacrylamide (9%) gel (Laemmli, 1970). Products were visualized by autoradiography.

Results

Gel Retardation Analysis of Fractions from DNA–Sepharose Chromatography. Two main peaks of [^{3}H]TCDD binding were observed following DNA–Sepharose chromarography of the reconstituted 33.3% ammonium sulfate pellet prepared from rat hepatic cytosol (Figure 1A). As previously noted (Henry et al., 1989), most (>90%) of the radioactivity contained in the flow-through fractions represents nonspecific [^{3}H]TCDD binding, while that eluting in the NaCl gradient is (>90%) specific binding. The peak eluting at approximately 0.19 M NaCl (peak 1) has been characterized by hydrodynamic and electrophoretic analyses as the monomeric, 100-kDa subunit of the Ah receptor (Henry et al., 1989; Poland & Glover, 1988; see also Figure 5), whereas that eluting at approximately 0.28 M NaCl (peak 2) is a multimeric, 160–200-kDa form (Henry et al., 1989). The latter also has physicochemical characteristics similar to those of the receptor extracted from nuclei of [^{3}H]TCDD-treated animals (Gasiewicz & Bauman, 1987; Prokipcak & Okey, 1988; Henry et al., 1989), suggesting that it is the transcriptionally active form of the receptor.

Gel retardation analysis of crude rat hepatic cytosol incubated with or without [^{3}H]TCDD revealed the presence of a TCDD-dependent protein–DNA complex (Figure 1B; compare lanes 1 and 2, arrow). These results are similar to those reported by other investigators (Fujisawa-Sehara et al., 1988; Nemoto et al., 1990). Analyses of fractions from DNA–Sepharose chromatography indicated that TCDD-dependent binding to the oligonucleotide containing the recognition motif coeluted with the [^{3}H]TCDD contained in peak 2 (compare Figure 1B, lanes 4–26, with Figure 1A). Only extremely weak bands migrating in the same position as the TCDD-dependent bands seen in crude cytosol were observed when fractions from a similar preparation incubated in the absence of TCDD were analyzed (not shown); thus these bands are TCDD-dependent. The flow-through fractions also contained TCDD-dependent protein–DNA complexes (Figure 1B, lane 3). The inability of this material to bind to the DNA–Sepharose column was thought to be a consequence of either (1) its ability to interact with other proteins and RNA contained in cytosol, (2) the presence of free receptor in equilibrium with DNA-bound receptor, or (3) a failure to reach equilibrium binding during chromatography [see Henry et al. (1989)]. This was further suggested by an experiment in which both DNA-binding peaks 1 and 2 were observed following rechromatography of aliquots of the flow-through fractions (not shown). For the remainder of these studies, peak 2 from DNA–Sepharose chromatography was used as the source for the partially purified, DNA-binding form of the receptor.

Competition experiments revealed that the peak 2 material had the DNA sequence specificity expected for the liganded receptor. Fractions equivalent to 57–72 (Figure 1A) were pooled and concentrated by ammonium sulfate precipitation. Using gel retardation, we found that the unlabeled wild-type recognition sequence competed well with the ^{32}P-labeled oligonucleotide for the liganded receptor (Figure 2, lanes 2–4). In contrast, an oligonucleotide containing methylated cytosine residues in the recognition sequence, which binds to the receptor weakly (Shen & Whitlock, 1989), did not compete

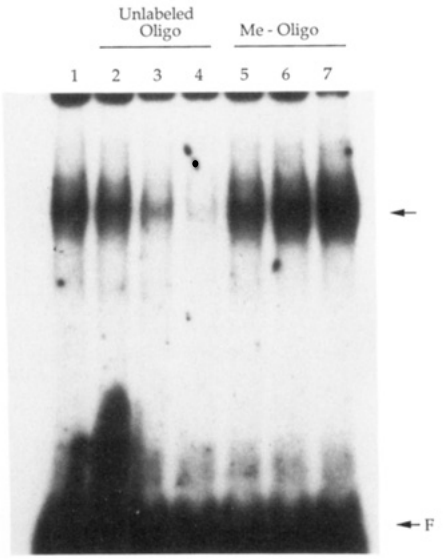

FIGURE 2: DNA sequence specificity of peak 2. Rat hepatic cytosol was chromatographed on DNA–Sepharose, and fractions equivalent to 57–72 in Figure 1A were pooled and concentrated by ammonium sulfate precipitation. This preparation was analyzed by gel retardation in the absence (lane 1) or the presence of a 1-, 5-, or 10-fold excess of unlabeled oligonucleotide (unlabeled oligo, lanes 2–4) or the presence a of 1-, 5-, or 10-fold excess of unlabeled methylated oligonucleotide (Me-oligo, lanes 5–7). The arrow indicates TCDD-dependent bands. F = free ^{32}P-labeled oligonucleotide.

(Figure 2, lanes 5–7). These observations indicated that the 160–200-kDa (peak 2) form of the receptor specifically binds to its cognate DNA recognition sequence.

Evidence That the Monomeric Ligand-Binding Form of the Receptor Alone Does Not Bind to the Recognition Motif. We sought to determine whether the 100-kDa ligand-binding form of the Ah receptor (peak 1) could bind to the DNA recognition motif. These experiments were performed by using peak 1 prepared from guinea pig and hamster hepatic cytosols because we can isolate substantial amounts of both peaks 1 and 2 from these species, whereas only small amounts of peak 1 are observed in rat hepatic cytosol under any incubation conditions (Figure 1A; Henry et al., 1989).

Gel retardation analyses (Figure 3A) indicated that crude cytosol preparations from guinea pig and hamster liver contained a TCDD-dependent DNA-binding protein similar to that observed in the rat. The analyses of peak 2 from both guinea pig and hamster produced these same retarded bands (Figure 3B). The finding that these bands were competitively inhibited by the unlabeled wild-type sequence but not by the methylated sequence (Figure 3B, lanes 3,4) indicated that they had the DNA sequence specificity expected for the liganded Ah receptor.[2] In contrast to peak 2, only faint bands were observed when peak 1 material, which contained an equivalent amount of TCDD specific binding, was analyzed by gel retardation (Figure 3B, lane 1). The presence of these faint bands is likely due to a small amount of peak 2 material present in the peak 1 fractions [see Henry et al. (1989)]. These results strongly suggest that the 100-kDa form of the receptor (peak 1) alone does not bind to the recognition motif and further imply that the 160–200-kDa form of the Ah receptor contains an additional component which confers the ability

[2] Note that a TCDD-dependent and recognition sequence specific doublet was observed by using both cytosol and partially purified peak 2 from guinea pig liver. Work is in progress to determine whether this is due to an association of the receptor–DNA complex with an additional protein present in guinea pig cytosol.

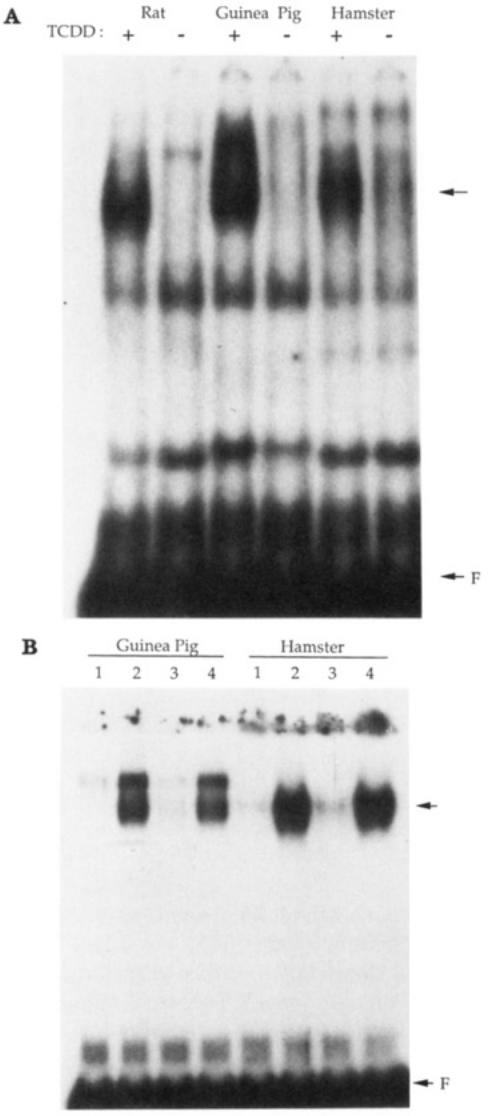

FIGURE 3: Gel retardation analysis of crude cytosol and partially purified receptor forms from different species. (A) Hepatic cytosol (~15 mg of protein/mL) from rat (lanes 1 and 2), guinea pig (lanes 3 and 4), or hamster (lanes 5 and 6) was incubated (2 h, 20 °C) with (+) or without (–) TCDD. Aliquots were tested for their interaction with the ^{32}P-labeled oligonucleotide by using gel retardation. (B) Guinea pig and hamster hepatic cytosols were incubated with [^{3}H]TCDD and chromatographed on DNA–Sepharose; the fraction containing maximum radioactivity representing peak 1 (~0.17 M NaCl) (lane 1) or peak 2 (~0.28 M NaCl) (lanes 2–4) was supplemented with 4 mg of BSA/mL and analyzed by gel retardation. In the case of peak 2, competition for binding to the ^{32}P-labeled oligonucleotide was tested by using a 10-fold excess of unlabeled oligonucleotide (lane 3) or methylated oligonucleotide (lane 4).

to bind to its DNA recognition motif.

Protein–DNA Cross-Linking Studies. We used a protein–DNA cross-linking technique to determine how many subunits bind to the recognition motif and the molecular masses of these individual subunits. The incorporation of BrdU into DNA in place of deoxythymidine facilitates the induction of protein–DNA cross-linking by UV irradiation (Lin & Riggs, 1974). A double-stranded oligonucleotide containing three BrdU residues within the recognition sequence and a fourth near one end was generated as described under Experimental Procedures. Digestion with the restriction enzyme *Msp*I removes the latter residue but does not detectably alter the profile of the cross-linked receptor–DNA complexes following SDS gel electrophoresis (not shown). Therefore, we assume that the BrdU residue near the end of the oligonucleotide did not

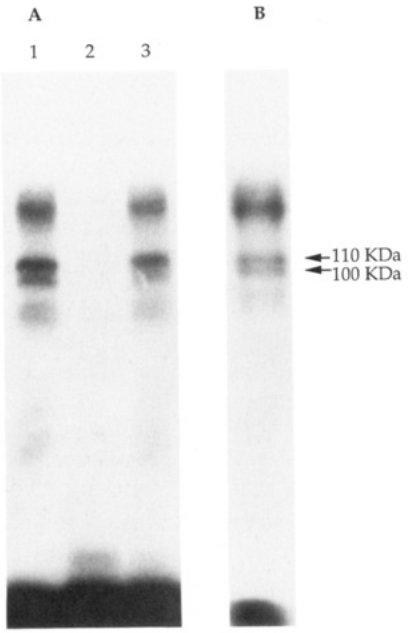

FIGURE 4: Cross-linking of partially purified receptor to DNA. Peak 2 was partially purified from rat hepatic cytosol as described in Figure 2. (A) Aliquots were incubated with the ^{32}P-labeled, BrdU-substituted oligonucleotide in the absence (lane 1) or presence of a 5-fold excess of unlabeled oligonucleotide (lane 2) or in the presence of a 5-fold excess of methylated oligonucleotide (lane 3). The samples were UV-irradiated and analyzed by SDS–polyacrylamide gel electrophoresis. (B) Peak 2 was incubated with the ^{32}P-labeled, BrdU-substituted oligonucleotide, UV irradiated for 15 min at 0 °C, and the receptor–DNA complexes were fractionated by gel electrophoresis. The position of the retarded band was determined by autoradiography. The band was excised as a gel slice, and the cross-linked products were analyzed by SDS–polyacrylamide gel electrophoresis and autoradiography.

contribute to the results presented below.

SDS–polyacrylamide gel analysis of peak 2 cross-linked to BrdU-substituted DNA showed the presence of three cross-linked protein–DNA complexes migrating at approximate molecular masses of 100, 110, and 220 kDa (Figure 4A, lane 1). The formation of cross-linked complexes was dependent on both UV irradiation and the presence of BrdU residues in the DNA (not shown). Competition experiments using the unlabeled wild-type (lane 2) and methylated (lane 3) sequences indicated that the three cross-linked complexes have the DNA sequence specificity characteristic of the liganded Ah receptor. Other studies using receptor contained in crude cytosol showed that DNase I digestion of the receptor–DNA complexes prior to electrophoresis produced essentially no change in the mobilities of these complexes on SDS–polyacrylamide gels (Elferink et al., 1990). Thus, cross-linking of the receptor subunits to the oligonucleotide had no apparent effect on the determined molecular weights of the radiolabeled protein bands. To further determine whether the three cross-linked protein–DNA complexes reflected the binding of receptor subunits to DNA, we identified the position of the cross-linked, TCDD-dependent protein–DNA complex in a retardation gel. We then isolated the complex from the gel and subjected it to SDS–polyacrylamide gel electrophoresis. Autoradiography of this gel revealed the presence of the same three cross-linked protein–DNA complexes (Figure 4B).

The simplest interpretation of these studies is that the DNA-binding form of the receptor is composed of at least three different protein subunits, all of which bind to DNA. However, an alternative interpretation, which we favor, is that the 220-kDa band represents the cross-linking of the two smaller (100 and 110 kDa) proteins to a single DNA molecule. Hydrodynamic studies suggested that the molecular mass of the DNA-binding form of the receptor is in the 160–200-kDa range (Prokipcak & Okey, 1988; Henry et al., 1989). Furthermore, the time course of the cross-linking reaction using crude cytosol indicated that the two smaller protein–DNA complexes form prior to the formation of the larger complex (Elferink et al., 1990). This suggests that the generation of the largest complex requires more than one cross-linking event. Therefore, we interpret the above results to indicate that the DNA-binding form of the receptor is composed of at least two protein subunits of 100 and 110 kDa.

Protein–Protein Cross-Linking Studies. Chemical cross-linking reagents have been used to examine the subunit composition of a variety of proteins. In particular, the bisimidates specifically react with lysine residues under mild conditions, and they are hydrolyzed rapidly, thus reducing the likelihood of random linkages (Aranyi, 1984; Aranyi et al., 1984). Using DMP (effective reagent length 0.73 nm), we sought to establish whether the 100-kDa ligand-binding form of the Ah receptor (peak 1) associates with another protein or proteins to generate the receptor form (peak 2) with high DNA-binding affinity. Initial experiments were performed with partially purified peaks 1 and 2 from guinea pig hepatic cytosol for the reasons described above.

Analysis of [^{3}H]TCDD photoaffinity-labeled peak 1 from guinea pig with or without DMP cross-linking revealed only a single peak of radioactivity migrating at approximately 100 kDa (Figure 5A). DMP appeared to produce some (approximately 30%) degradation of this form of the receptor as indicated by decreased radioactivity in the 100-kDa peak and the increase at the small molecular mass region of the gel. Thus, we were unable to cross-link the 100-kDa (peak 1) form of the receptor into a higher molecular mass species. These results are consistent with the idea that this form is monomeric and is the ligand-binding polypeptide.

In the absence of DMP, the majority of the radioactivity associated with peak 2 also migrated at approximately 100 kDa (Figure 5B). In the presence of DMP, there was a partial shift of radioactivity from the 100-kDa peak to a slightly broader peak at approximately 210 kDa. Neither radioactive peak was observed when these incubations were performed by using peak 2 prepared in the presence of [^{3}H]TCDD plus a 100-fold excess of the competing ligand TCDF (not shown); thus, the binding observed is specific binding. These studies indicate that the 100-kDa form of the receptor in peak 2 can be cross-linked into a higher molecular mass species. This finding implies that the receptor in peak 2 isolated from guinea pig liver exists as a complex of a 100-kDa ligand-binding protein (peak 1) and an additional protein or proteins of approximately 110 kDa.

Analysis of [^{3}H]TCDD photoaffinity-labeled peak 2 from rat liver showed results similar to those found in guinea pig hepatic cytosol except that the cross-linked receptor migrated at approximately 195 kDa (not shown). Figure 5C shows the result of an experiment in which [^{125}I]IBr_2DD photoaffinity-labeled peak 2 from rat hepatic cytosol was analyzed. Again, in the absence of DMP, the major band of radioactivity migrated at approximately 100 kDa. In the presence of DMP, a band was also observed at approximately 190 kDa and several fainter bands were observed at higher molecular masses; the latter may represent multiple cross-linked conformations. Combining individual results obtained with [^{3}H]TCDD and [^{125}I]IBr_2DD, the molecular masses of the major bands of radioactivity observed in rat hepatic cytosol

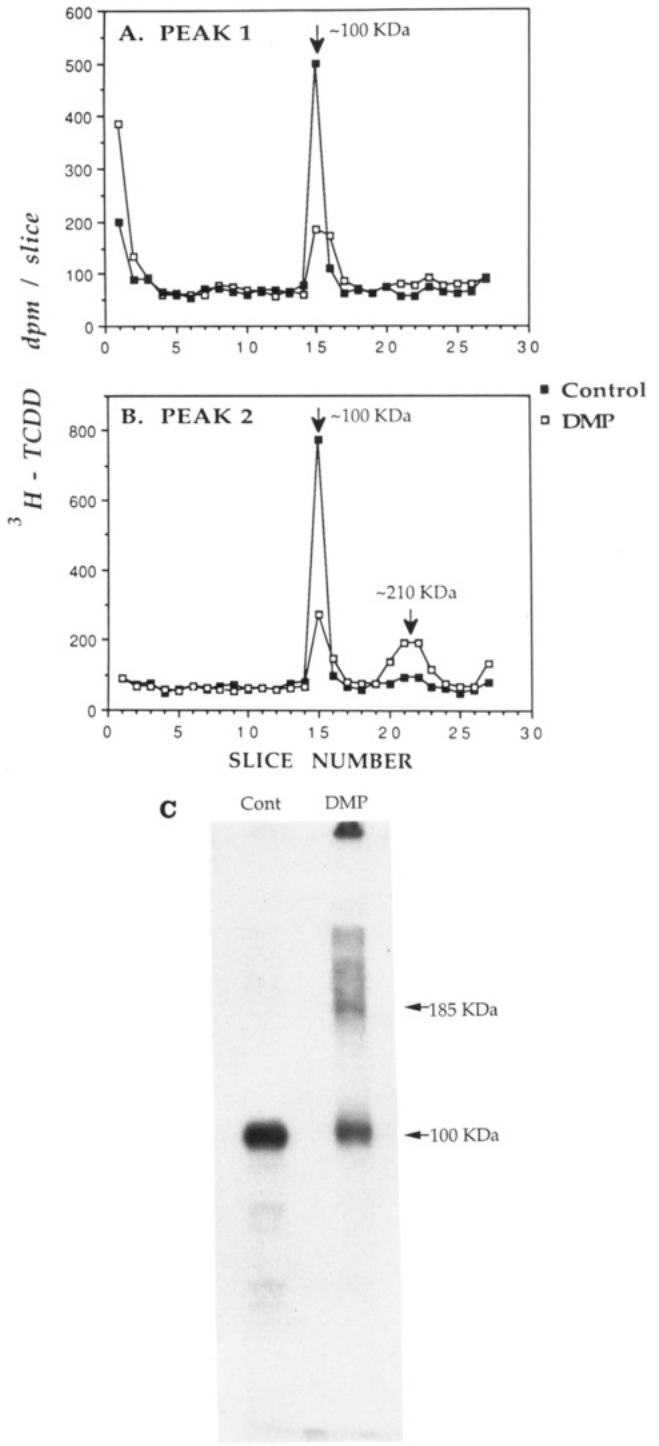

FIGURE 5: Protein–protein cross-linking of receptor peaks 1 and 2. Guinea pig hepatic cytosol was incubated with [^{3}H]TCDD and fractionated by DNA–Sepharose chromatography to obtain pooled samples of peaks 1 (A) and 2 (B) (see Figure 3). Aliquots were treated with DMP or vehicle (control) and subjected to UV irradiation to covalently bind [^{3}H]TCDD to receptor. Protein was precipitated and dissolved in sample buffer for SDS–polyacrylamide gel electrophoresis. After electrophoresis, lanes were cut into 5-mm slices for radioactivity determination. (C) Peak 2 was obtained as above from rat hepatic cytosol after incubation with [^{125}I]IBr_2DD. Untreated (Cont) and DMP-treated (DMP) samples were cross-linked and photoaffinity-labeled, and electrophoresis was performed as in (A) and (B). Labeled bands were visualized by autoradiography.

in the presence of DMP were 98 ± 5 and 191 ± 8 kDa (n = 4). Thus, peak 2 from rat hepatic cytosol is a high-affinity DNA-binding complex of approximately 200 kDa made up of a ligand-binding polypeptide of approximately 100 kDa and an additional protein(s).[3] These results combined with the protein–DNA cross-linking studies also suggest that the two proteins interact independently of and prior to DNA binding.

Discussion

We have previously distinguished and characterized three forms of the Ah receptor present in crude hepatic cytosol [ligand unoccupied, occupied monomer (peak 1), and occupied multimeric (peak 2)] on the basis of their hydrodynamic properties, overall surface charge, molecular mass, and binding to calf thymus DNA (Gasiewicz & Bauman, 1987; Henry et al., 1989). In this report, we have extended these observations by demonstrating that the form which has the highest affinity for calf thymus DNA also binds to the DNA recognition motif contained within the enhancer element located upstream of the structural gene for CYPIA1. Although additional components may be necessary for the regulation of this gene in vivo, our results imply that crude cytosol contains the constituents necessary for the ligand-induced transformation of the receptor to a form which recognizes a specific DNA sequence. Thus, cytosol provides a useful model system for analyzing the molecular changes that occur in the receptor following ligand binding.

We also present evidence consistent with our model (Henry et al., 1989) that the form of the receptor which binds to the DNA recognition motif exists as a heteromeric structure. This evidence is based on (1) the characterization of peaks 1 and 2 as monomeric and multimeric ligand-binding proteins, respectively, (2) the differing ability of partially purified peaks 1 and 2 to bind to the recognition motif in a sequence-specific fashion, (3) the determination by protein–DNA cross-linking that in rat hepatic cytosol two polypeptides of approximately 100 kDa and 110 kDa bind to DNA in a TCDD-dependent and sequence-specific manner, (4) the ability of the 100-kDa subunit to cross-link with another protein or proteins contained in cytosol, but not with itself, and (5) the finding that only the 100-kDa subunit is associated with ligand. Furthermore, these data are consistent with other lines of evidence suggesting that the DNA-binding form of the Ah receptor may contain more than one subunit. First, both the cytosolic receptor form (peak 2) that has the highest affinity for calf thymus DNA and that extracted from nuclei of TCDD-treated animals have been shown by hydrodynamic analysis to be of greater M_r than the ligand-binding monomeric subunit (Prokipcak & Okey, 1988; Henry et al., 1989). Second, our previous studies have suggested that the process of receptor transformation involves a non-ligand-binding factor present in cytosol (Henry et al., 1989). Third, complementation studies examining the phenotype of cell lines demonstrating varying responsiveness to TCDD induction of CYPIA1 imply that at least three genes contribute to receptor function (Hankinson, 1983; Miller et al., 1983; Hankinson et al., 1985; Karenlampi et al., 1988). While our present studies imply the direct association of only one component with the ligand-binding subunit (100 kDa) contained in rat hepatic cytosol, they do not rule out the possibility that other species- and tissue-specific factors may modulate the function of the heterodimer we have identified.

Recently, some evidence was presented suggesting that a 4–5S species of the Ah receptor is the form that binds to the

[3] Photoaffinity-labeling studies were also performed by using crude rat hepatic cytosol under the same conditions as used for the DNA cross-linking studies with the exception that unlabeled BrdU-substituted DNA was used. Only the 100-kDa band on SDS gels showed the presence of radiolabeled ligand (Elferink et al., 1990), confirming that only one of the two protein subunits that cross-link to DNA contains ligand.

DNA recognition sequence (Nemoto et al., 1990). This group found that the salt-induced DNA-binding activity of the receptor prepared from rat hepatic cytosol coeluted in the 4–5S region of glycerol gradients. In our hands, the partially purified monomeric ligand-binding form (peak 1) of the receptor sediments in sucrose gradients as a 4–5S species, and the multimeric form (peak 2) sediments as a 7–8S species (Henry et al., 1989). Although our results appear to disagree with those described by Nemoto et al. (1990), it is possible that the DNA-binding activity detected by this group may also have resulted from the interaction of the monomeric ligand-binding subunit (100 kDa) with the 110-kDa protein we have presently identified. Due to their similar molecular mass, both proteins, if dissociated, may cosediment in glycerol gradients. Thus, the receptor subunits may have dissociated and cosedimented during centrifugation and then reassociated under the conditions used by Nemoto et al. (1990) in the DNA-binding assay. Notably, the glucocorticoid receptor homodimer has been shown to dissociate during glycerol gradient centrifugation (Eriksson & Wrange, 1990).

Our data are compatible with a model in which the DNA-binding form of the Ah receptor is composed of at least two different subunits, both of which interact with DNA and only one of which is ligand-bound. This model is also consistent with the recent report that the molar ratio of [^{125}I]IBr$_2$DD to ^{32}P-labeled recognition sequence in the Ah receptor–DNA complex is about 1 (Denison et al., 1989). In contrast, both the estrogen and glucocorticoid receptors bind to their DNA recognition sequences as homodimers (Kumar & Chambon, 1988; Tsai et al., 1988). This is consistent with the palindromic structure of their cognate DNA recognition motifs (Beato, 1989). However, the DNA recognition sequence for the Ah receptor lacks symmetry (Denison et al., 1989). In addition, the Ah receptor-defective variants fall into several complementation groups (Miller et al., 1983; Hankinson et al., 1985; Karenlampi et al., 1988), whereas the analogous glucocorticoid receptor-defective cells are in a single complementation group (Yamamoto, 1985). Together these observations suggest that the Ah receptor is not a member of the class of transcription factors represented by the steroid hormone receptors.

The finding that the DNA-binding form of the Ah receptor is composed of (at least) two different polypeptides imposes constraints upon techniques to be used for purifying the functional form of this receptor. Until now, many experimental approaches have been based on the assumption that a single ligand-binding polypeptide or a homodimer of this polypeptide comprised the functional DNA-binding complex. Our present and previous (Henry et al., 1989) data indicate that future purification strategies need to consider the existence of non-DNA-binding forms and to maintain the association of the different polypeptides which constitute the functional form of the receptor.

On the basis of the results presented here, it appears that the ligand-bound monomeric 100-kDa subunit must interact with at least one other protein of approximately 110 kDa to generate a receptor that specifically recognizes the enhancer elements upstream of the CYPIA1 gene. Although the exact ligand-dependent structural change that allows the 100-kDa subunit to recognize this additional protein is not known, it has been hypothesized that this initial step involves dissociation of hsp 90 (Perdew, 1988; Wilhelmsson et al., 1990). We speculate that this dissociation allows a domain on the ligand-binding protein to specifically interact with the additional protein. The formation of the heterodimer results in a structural conformation that has both high affinity and specificity for the core sequence of the enhancer element.

The steroid hormone receptors are known to interact with DNA by homodimerization and the formation of a zinc-finger structural motif. However, several other structural motifs including helix–turn–helix and that formed by a leucine-zipper protein–protein dimerization have been described [reviewed by Johnson and McKnight (1989)]. For example, the heterodimerization of c-jun and c-fos oncoproteins via a leucine repeat helix results in an apparent conformational change of both polypeptides such that each interacts with gene promoters containing AP-1 binding sites (Halazonetis et al., 1988; Nakabeppu et al., 1988). Recently, the transcription factor PTFI has also been identified as a heterodimeric structure in which both proteins bind simultaneously to the same DNA molecule (Roux et al., 1989). The exact structural conformation that enables both subunits of the Ah receptor to interact with DNA in a sequence-specific manner remains to be determined. However, we envision that the regulated interaction of these two subunits, in possible combination with other polypeptides, may play a role in the species- and tissue-specific gene expression observed following exposure of animals to TCDD.

ACKNOWLEDGMENTS

We are indebted to George Rucci and Donna Galeazzi for excellent technical assistance and to Dr. Emily Shen for preparing the methylated oligonucleotide sequence. We thank Dr. James P. Whitlock, Jr., for helpful suggestions throughout the course of this research and during the preparation of the manuscript.

REFERENCES

Aranyi, P. (1984) *Biochem. Biophys. Res. Commun. 119*, 64–68.

Aranyi, P., Radanyi, C., Renoir, M., Devin, U., & Baulieu, E.-E. (1988) *Biochemistry 27*, 1330–1336.

Beato, M. (1989) *Cell 56*, 335–344.

Cuthill, S., Poellinger, L., & Gustafsson, J.-A. (1987) *J. Biol. Chem. 262*, 3477–3481.

Denis, M., Poellinger, L., Wikstrom, A.-C., & Gustafsson, J.-A. (1988) *Biochem. Biophys. Res. Commun. 155*, 801–807.

Denison, M. S., Fisher, J. M., & Whitlock, J. P., Jr. (1988) *J. Biol. Chem. 263*, 17221–17224.

Denison, M. S., Fisher, J. M., & Whitlock, J. P., Jr. (1989) *J. Biol. Chem. 264*, 16478–16482.

Elferink, C. J., Gasiewicz, T. A., & Whitlock, J. P., Jr. (1990) *J. Biol. Chem. 265*, 20708–20712.

Eriksson, P., & Wrange, O. (1990) *J. Biol. Chem. 265*, 3535–3542.

Fisher, J. M., Wu, L., Denison, M. S., & Whitlock, J. P., Jr. (1990) *J. Biol. Chem. 265*, 9676–9681.

Fujisawa-Sehara, A., Sogawa, K., Nishi, C., & Fujii-Kuriyama, Y. (1986) *Nucleic Acids Res. 14*, 1465–1577.

Fujisawa-Sehara, A., Yuamane, M., & Fujii-Kuryama, Y. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5859–5863.

Gasiewicz, T. A., & Neal, R. A. (1982) *Anal. Biochem. 124*, 1–11.

Gasiewicz, T. A., & Bauman, P. A. (1987) *J. Biol. Chem. 262*, 2116–2120.

Gonzalez, F. J., & Nebert, D. W. (1985) *Nucleic Acids Res. 13*, 7269–7288.

Gustafsson, J.-A., Carlstedt-Duke, J., Poellinger, L., Okret, S., Wikstrom, A.-C., et al. (1987) *Endocr. Rev. 8*, 185–234.

Halazonetis, T. D., Georgopoulos, K., Greenberg, M. E., & Leder, P. (1988) *Cell 55*, 917–924.

Hankinson, O. (1983) *Somatic Cell Genet. 9*, 497-514.
Hankinson, O., Anderson, R. D., Birren, B. W., Sander, F., Negishi, M., & Nebert, D. W. (1985) *J. Biol. Chem. 260*, 1790-1795.
Hannah, R. R., Lund, J., Poellinger, L., Gillner, M., & Gustafsson, J.-A. (1986) *Eur. J. Biochem. 156*, 237-242.
Henry, E. C., Rucci, G., & Gasiewicz, T. A. (1989) *Biochemistry 28*, 6430-6440.
Johnson, P. F., & McKnight, S. L. (1989) *Annu. Rev. Biochem. 58*, 799-839.
Jones, P. B. C., Galeazzi, D. R., Fisher, J. M., & Whitlock, J. P., Jr. (1985) *Science 227*, 1499-1502.
Karenlampi S. O., Legraverend, C., Gudas, J. M., Carremanzana, N., & Hankinson, O. (1988) *J. Biol. Chem. 263*, 10111-10117.
Kumar, V., & Chambon, P. (1988) *Cell 55*, 145-156.
Laemmli, U. K. (1970) *Nature 227*, 680-685.
Landers, J. P., Piskorska-Pliszczynska, J., Zacharewski, T., Bunce, N. J., & Safe, S. (1989) *J. Biol. Chem. 264*, 18463-18471.
Lin, S.-Y., & Riggs, A. D. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 947-951.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Miller, A. G., Israel, D. I. & Whitlock, J. P., Jr. (1983) *J. Biol. Chem. 258*, 3523-3527.
Nakabeppu, Y., Ryder, K., & Nathans, D. (1988) *Cell 55*, 907-915.
Nemoto, T., Mason, G. G. F., Wilhelmsson, A., Cuthill, S., Hapgood, J., Gustafsson, J.-A., & Poellinger, L. (1990) *J. Biol. Chem. 265*, 2269-2277.
Perdew, G. H. (1988) *J. Biol. Chem. 263*, 13082-13805.
Poland, A., & Knutson, J. C. (1982) *Annu. Rev. Pharmacol. Toxicol. 2*, 517-554.
Poland, A., & Glover, E. (1987) *Biochem. Biophys. Res. Commun. 146*, 1439-1449.
Poland, A., & Glover, E. (1988) *Arch. Biochem. Biophys. 261*, 103-111.
Poland, A., Glover, E., Ebetino, F. H., & Kende, A. S. (1986) *J. Biol. Chem. 261*, 6352-6365.
Prokipcak, R. D., & Okey, A. B. (1988) *Arch. Biochem. Biophys. 267*, 811-828.
Roux, E., Strubin, M., Hagenbuchle, O., & Wellauer, P. K. (1989) *Genes Dev. 3*, 1613-1624.
Shen, E. S., & Whitlock, J. P., Jr. (1989) *J. Biol. Chem. 264*, 17754-17758.
Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlmann, K., Gustafsson, J.-A., Tsai, M. H., & O'Malley, B. W. (1988) *Cell 55*, 361-369.
Whitlock, J. P., Jr. (1990) *Annu. Rev. Pharmacol. Toxicol. 30*, 251-277.
Wilhelmsson, A., Wikstrom, A.-C., & Poellinger, L. (1986) *J. Biol. Chem. 261*, 13456-13463.
Wilhelmsson, A., Cuthill, S., Denis, M., Wikstrom, A.-C., Gustafsson, J.-A., & Poellinger, L. (1990) *EMBO J. 9*, 69-76.
Winter, R. B., & von Hipple, P. H. (1981) *Biochemistry 20*, 6948-6960.
Yamamoto, K. R. (1985) *Annu. Rev. Genet. 19*, 209-252.

Biosynthesis of the *Escherichia coli* Siderophore Enterobactin: Sequence of the *entF* Gene, Expression and Purification of EntF, and Analysis of Covalent Phosphopantetheine[†,‡]

Frank Rusnak, Masahiro Sakaitani, Dale Drueckhammer,[§] Janice Reichert, and Christopher T. Walsh*
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115
Received September 11, 1990; Revised Manuscript Received December 12, 1990

ABSTRACT: The sequence of the *entF* gene which codes for the serine activating enzyme in enterobactin biosynthesis is reported. The gene encodes a protein with a calculated molecular weight of 142006 and shares homologies with the small subunits of gramicidin S synthetase and tyrocidine synthetase. We have subcloned and overexpressed *entF* in a multicopy plasmid and attempted to demonstrate L-serine-dependent ATP-[^{32}P]PP$_i$ exchange activity and its participation in enterobactin biosynthesis, but the overexpressed enzyme appears to be essentially inactive in crude extract. A partial purification of active EntF from wild-type *Escherichia coli*, however, has confirmed the expected activities of EntF. In a search for possible causes for the low level of activity of the overexpressed enzyme, we have discovered that EntF contains a covalently bound phosphopantetheine cofactor.

Under conditions of iron deprivation, many microorganisms synthesize and secrete low molecular weight compounds termed siderophores which bind ferric ions with high affinity and are used to supply iron for metabolic processes. *Escherichia coli* and other enteric bacteria synthesize the catechol-containing siderophore enterobactin, a macrocyclic trilactone comprised of three molecules each of 2,3-dihydroxybenzoate (2,3-DHB) and L-serine (Figure 1A). Much is now known about the intricate metabolic machinery of *E. coli* involved in enterobactin biosynthesis, uptake, and processing. Thus, the *ent* genes encode the biosynthetic enzymes (Young et al., 1971; Luke & Gibson, 1971; Woodrow et al., 1975a,b), the *fep* genes are required for the uptake of the ferric enterobactin complex (Pierce et al., 1983; Pierce & Earhart, 1986;

[†] Supported by NIH Grant GM20011 (C.T.W.) and NIH Postdoctoral Fellowships GM12806 (F.R.), GM12014 (D.D.), and GM13804 (J.R.).
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05325.
* Author to whom correspondence should be addressed.
[§] Present address: Department of Chemistry, Stanford University, Stanford, CA 94305.

0006-2960/91/0430-2916$02.50/0