Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 11503–11507, December 1992
Biochemistry

DNA binding and heteromerization of the *Drosophila* transcription factor chorion factor 1/ultraspiracle

ANASTASIA M. KHOURY CHRISTIANSON*†, DENNIS L. KING*, EVIE HATZIVASSILIOU‡, JORGE E. CASAS*, PAUL L. HALLENBECK§, VERA M. NIKODEM§, S. ALEX MITSIALIS¶, AND F. C. KAFATOS*‖**

*Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138; ‡Department of Biochemistry and ¶Evans Department of Clinical Research, The University Hospital and Department of Medicine, Boston University School of Medicine, Boston, MA 02118; §Genetics and Biochemistry Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892; and ‖Institute of Molecular Biology and Biotechnology, P.O. Box 1527, Heraklion 711 10, Crete, Greece

Contributed by F. C. Kafatos, August 19, 1992

ABSTRACT **The *Drosophila* chorion factor 1/ultraspiracle (CF1/USP) transcription factor, a homologue of the retinoid X receptor, is a developmentally important member of the family of nuclear (steroid) hormone receptors. Using newly developed monoclonal antibodies and a full-length bacterially produced protein, we have studied in detail the *in vitro* DNA-binding properties of this factor and aspects of its distribution *in vivo*. During oogenesis, CF1/USP is present both in germline cells and in the somatic follicular epithelium. We have determined the optimal binding site of partially purified bacterially produced CF1/USP by an *in vitro* selection procedure and also have characterized its binding to the follicular-specific chorion *s15* promoter. *In vitro* this bacterially produced factor is unusual in binding to a single element ("half-site"); simultaneous but noncoordinate binding to a second half-site is possible if these repeated elements are organized in direct orientation and spaced adequately. However, the factor interacts synergistically with several other nuclear hormone receptors: notably, it can form *in vitro* heteromers with mammalian thyroid and retinoic acid receptors, binding to two half-sites that are organized in either direct or inverted orientation. *In vivo* the factor most probably functions as a heterodimer, but its partner(s) remains to be determined.**

Steroid and other nuclear hormone receptors are a family of ligand-modulated transcription factors that regulate cell differentiation and development as well as homeostasis and reproduction (1, 2). These receptors bind to DNA promoter or enhancer motifs ("hormone response elements" or HRE), which typically contain two copies of short, receptor-specific sequences called "half-sites," oriented either as direct (e.g., refs. 3–5) or as inverted (e.g., refs. 6 and 7) repeats. Although some members of the family bind to HREs as homodimers, others can heterodimerize, potentially enhancing the range of regulatory processes that can be served by a single HRE (8).

In *Drosophila melanogaster*, the paradigm for metazoan genetics, nine members of this family have been identified so far, with strategic roles in embryogenesis, postembryonic development, and morphogenesis (2). We first cloned one of these factors by its binding to a potential HRE in the chorion *s15* promoter and named it chorion factor 1 (CF1) (9); this receptor was also cloned independently by homology to the DNA-binding domain of vertebrate hormone receptors (10, 11). This orphan receptor, CF1/USP, is encoded by the ultraspiracle locus (*usp*) (10), a maternally and zygotically required gene with multiple functions in development (12). CF1/USP is most closely related to the subfamily of vertebrate retinoid X receptors (RXRs), suggesting that its ligand might be a retinoid metabolite (13–15).

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.

In our initial work (9), experiments using the chorion *s15* promoter and a bacterially produced protein fragment indicated that CF1/USP binds to a single half-site. We have now confirmed this unusual feature in detailed studies of the binding properties of this factor *in vitro*. CF1/USP binds moderately strongly to a single HRE half-site; a second half-site on the same DNA can be used simultaneously but not cooperatively, provided it is present as a direct repeat with sufficient spacing. In contrast, the factor is prone to interact with other members of the family, and *in vivo* most likely it functions as heterodimers with one or more as yet unknown partners. Binding studies *in vitro* demonstrate that this fly factor can assemble on DNA with mammalian thyroid hormone receptor (THR) and retinoic acid receptor (RAR).

MATERIALS AND METHODS

Bacterial CF1/USP Preparations. Truncated CF1 (T-CF1) was prepared as described (9). For production of CF1, a genomic DNA *usp* fragment [a gift from A. Oro (Salk Institute)] (12) was subcloned into the phage T7–*Escherichia coli* expression system (16). Host *E. coli* BL21 (DE3) was used for protein production as described (17). CF1 was partially purified on a phosphocellulose column. Details will be described elsewhere.

To prepare glutathione *S*-transferase fusion protein (GST-CF1), pGST-CF1 was constructed by inserting a blunt-ended *Nde* I/*Eco*RI fragment of pCF1 containing *usp* into blunt-ended pGEX-2T plasmid digested with *Bam*HI. The fusion protein was purified as described (18).

Optimal Site Selection. The optimal binding site of CF1 was determined as described (19–21). A 60-base-pair (bp) oligonucleotide containing a completely random 20-bp region (CCGCTCGAGGATCCN$_{20}$GAATTCCCGGGAAG) was end-labeled to high specific activity with polynucleotide kinase and was used in electrophoretic mobility shift assay (EMSA) gels. Nontransformed bacterial extract was used in parallel as the negative control. The retarded complex observed by autoradiography was excised, and the DNA was recovered (21). A broad region of the gel slice was excised (10 mm × 30 mm) to permit recovery of higher order complexes, and the DNA included therein was extracted and amplified by PCR with primers flanking the 20-bp degenerate region. This DNA was subjected to additional cycles of binding/EMSA

Abbreviations: CF1, chorion factor 1; T-CF1, truncated CF1; *usp*, ultraspiracle locus; CF1/USP, product of *usp*; HRE, hormone response element; THR, thyroid hormone receptor; hTHR and rTHR, human and rat THR; RXR, retinoid X receptor; hRXR, human RXR; RAR, retinoic acid receptor; hRAR, human RAR; mAb, monoclonal antibodies; EMSA, electrophoretic mobility shift assay; GST, glutathione *S*-transferase.

†Present address: DNX Corporation, 303 B College Road East, Princeton, NJ 08540.
**To whom reprint requests should be addressed.

purification/PCR amplification under progressively more stringent conditions (21). DNA recovered from the gel after cycles 3 and 5 was subcloned into a plasmid, and independent clones were sequenced.

Antibody Detection of USP on EMSA Gels. Standard EMSA gels were run as described (9). The gel was electrotransferred onto stacked nitrocellulose and Zeta-Probe (Bio-Rad) membranes, allowing the protein to adhere to nitrocellulose while the DNA travels through to the Zeta-Probe membrane. The nitrocellulose membrane was probed by antibodies. Details of this method will be described elsewhere.

Preparation of Nuclear Receptors Expressed in Cell Lines. The insert of plasmid pCF1 was subcloned into the expression vector PCDNAI (Invitrogen, San Diego). Total extracts from COS cells transiently transfected by the resulting plasmid were prepared as described by Kumar and Chambon (6). Control extracts were prepared from cells transfected with vector containing the CF1 insert in the anti-sense orientation. Extracts from COS cells expressing human RARα (hRARα), human RXRα (hRXRα), and human THRβ (hTHRβ) proteins were a generous gift of J. Ladias (N.E. Deaconess Hospital). The human AII apolipoprotein promoter sequence used corresponds to the AII-J region (22) and was a gift from C. Cladaras (Boston University). Nuclear extracts containing high levels of rTHRα were prepared from recombinant baculovirus-infected Sf9 cells (23). The thyroid response element used is a 72-bp *Hin*dIII–*Eco*RI restriction fragment prepared from plasmid pTK1AM3 containing the rat malic enzyme response element (23).

Follicle Cell Staining. Ovarian follicles were rinsed in Ringer's solution and fixed [4% methanol-free formaldehyde (Polysciences)/160 mM KCl/40 mM NaCl/4 mM Na_3EGTA/5 mM spermidine/2 mM spermine/30 mM Pipes, pH 7.4] for 10 min at room temperature. After a 10-min wash in PBST (PBS containing 0.1% Triton X-100) they were incubated for 2 hr with either AB11 or DH9 monoclonal antibody (mAb) supernatant. Follicles were then washed for 15–30 min in PBS and incubated for 1–2 hr with fluorescein-labeled, goat anti-mouse antibodies (Jackson ImmunoResearch) diluted 1:200 with PBS. After a final 15- to 30-min wash in PBS, whole mounts of follicles were prepared. Laser scanning confocal microscopy was performed with the MRC 600 confocal head (Bio-Rad) mounted on a Zeiss Axioskop microscope.

RESULTS

Tools for Studying CF1/USP. *usp* clones were expressed in bacteria, yielding three CF1/USP preparations that henceforth will be abbreviated T-CF1, CF1, and GST-CF1 (Fig. 1*A*). T-CF1 was obtained by expression of the original clone (9), which encodes 280 amino acid residues centered on the DNA-binding domain. CF1 was obtained from a full-length clone that encodes 508 residues. GST-CF1 was identical except that it was fused at its NH_2 terminus with GST, permitting subsequent purification on glutathione-agarose beads. Immunoblot (Western blot) analysis using mouse mAbs to CF1 revealed that all three preparations were polydisperse as a result of partial degradation (Fig. 1*B*). CF1 fractions enriched in full-length product were obtained by gradient salt elution from a phosphocellulose column and were used in most of the subsequent DNA binding studies.

Comparative Western blot analysis permitted mapping of the epitopes for two different mAbs. The DH9 mAb recognizes the extreme NH_2 terminus of the CF1/USP protein: it does not bind to T-CF1, but, importantly, it binds to glutathione-selected GST-CF1 fragments of 28 kDa—merely 2 kDa larger than the vector-encoded glutathione *S*-transferase alone. The NH_2 terminus is nonconserved even among RXR-type proteins (15), making DH9 an ideal tool for CF1/USP detection. For example, DH9 detects a single 55-kDa band in Western blots of nuclear extracts from a variety of *D. melanogaster* cell types, including the follicular epithelium (Fig. 1*B*); in contrast, it does not recognize the corresponding component of *Drosophila virilis* or the moth *Bombyx mori*. Fig. 1*C* shows immunostained follicles (stage 13) that confirm both the specificity of DH9 for the *D. melanogaster* (as opposed to *D. virilis*) CF1/USP protein and the presence of that protein in the nuclei of the follicular epithelium. Staining also revealed the existence of the protein in germ-line nurse cells as well as follicle cells from all other stages of oogenesis (data not shown). Comparable experiments have shown that the AB11 mAb binds to the same 55-kDa CF1/USP protein; it recognizes an epitope that is shared by the *D. virilis* and *B. mori* homologues (data not shown) and maps in the region of the first zinc finger (Fig. 1*A*).

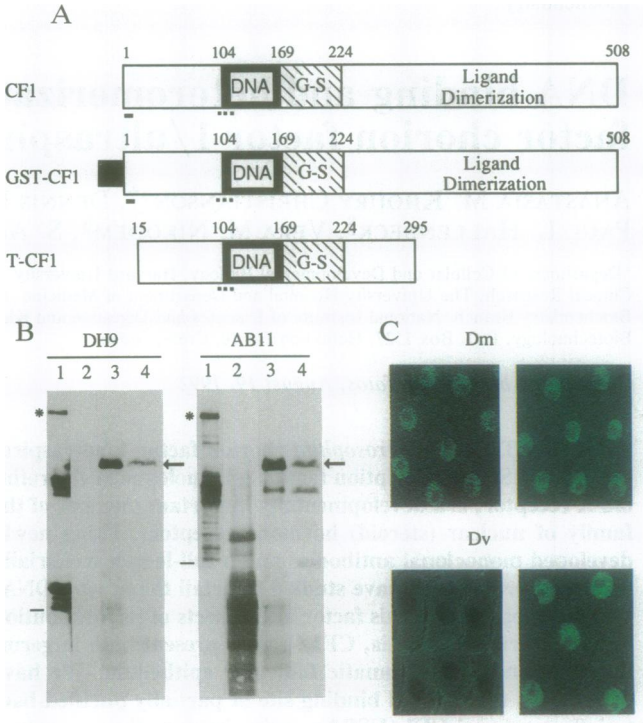

FIG. 1. (*A*) Schematic of the CF1/USP preparations. The zinc finger domain involved in DNA binding, a glycine- and serine-rich domain, and the ligand binding/dimerization domain are indicated. CF1 is the full-length bacterially produced CF1/USP protein (508 residues, ≈55 kDa), GST-CF1 is the fusion protein of the full-length CF1 with glutathione *S*-transferase protein (≈81 kDa), and T-CF1 is the original truncated CF1 fragment (9). The region of the DH9 epitope is shown by an underline, while the region of the AB11 epitope is signified by three dots. (*B*) Specificity of anti-CF1 mAb preparations. Bacterially expressed proteins GST-CF1 (lane 1), T-CF1 (lane 2), CF1 (lane 3), or *D. melanogaster* follicular nuclear extract (lane 4) were analyzed by Western blot with the indicated mAbs [DH9 (*Left*) or AB11 (*Right*)] as probes. The fusion protein GST-CF1 is indicated by an asterisk, and the full-length CF1, by an arrow. A dash indicates the lower size limit of detected degradation products. Note the absence of the DH9 epitope from T-CF1 (*Left*, lane 2). (*C*) Presence of CF1/USP in follicular epithelial nuclei of both *D. melanogaster* (Dm) (*Upper*) and *D. virilis* (Dv) (*Lower*). Stage 13 follicles were immunostained with mAb DH9 (*Left*) or mAb AB11 (*Right*). Note the presence of the protein in both species, as monitored by the AB11 mAb (*Upper Right* and *Lower Right*), the presence of the DH9 epitope in *D. melanogaster* (*Upper Left*), and its absence in *D. virilis* (*Lower Left*).

Binding to the Chorion *s15* Promoter. A 0.13-kbp fragment of the *D. melanogaster* chorion *s15* promoter was used for DNase I footprinting (DNase cleavage inhibition patterns) (9), methylation interference (method of refs. 21 and 24), and $KMnO_4$ hydroxylation interference (25) studies of this factor. Both the CF1 and T-CF1 footprints encompassed the up-

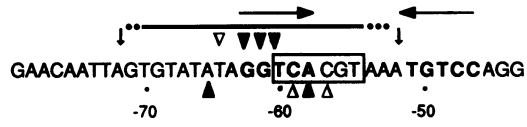

FIG. 2. Binding of CF1/USP to the *D. melanogaster* chorion *s15* promoter. The invariant chorion-specific hexamer is boxed. A region of imperfect dyad symmetry is indicated by horizontal arrows and boldface letters. The limits of CF1 binding as detected by DNase I protection analysis are indicated by an overline, with dots defining the maximum possible extent of the protected region. Bases important for binding as defined by methylation and hydroxylation interference analyses are indicated by solid (strong effect) or open (weak effect) arrowheads. Arrowheads above the sequence indicate contacts with the top strand (shown); those underneath indicate noncoding strand contacts. Note the absence of interference by modification of the downstream half-site. Vertical arrows indicate DNase I hypersensitive sites induced by CF1 binding.

stream but clearly not the downstream half-site of the imperfect palindromic *s15* HRE, GGTCACGTAAATGTCC. Fig. 2 summarizes the results. The footprint is bracketed by hypersensitive sites at promoter positions −72 and −52, the latter of which is located just upstream of the TGTCC half-site. Methylation or hydroxylation at each nucleotide position of the GGTCA half-site and at three neighboring positions interfered with binding; there was no interference from modification of the TGTCC half-site.

Optimal CF1 Binding Site. To characterize in more detail the nucleotide requirements for CF1/USP binding to DNA, we used a binding-site-selection technique (19–21) based on the polymerase chain reaction. The partially purified CF1 preparation was mixed with a 60-bp oligonucleotide bearing a central, completely random 20-bp region (see *Materials and Methods*). Thirty-eight sequences from cycle 5 were examined and showed remarkable agreement, indicating very selective binding to a single half-site. Fig. 3 shows the alignment of the sequences and a table summarizing the tolerance for each base at each position. The optimal CF1 binding sequence was the nonamer GGGGTCACG, which encompasses the GGTCA half-site but differs in the first two positions from the corresponding nonamer of the *s15* promoter (TAGGTCACG). To verify this preference, an oligonucleotide encompassing the region of the *s15* binding site (−36 to −84) and a variant (same size oligonucleotide) bearing the alternative nonamer were tested in parallel in the EMSA procedure with progressively lower protein concentration. Fig. 3 confirms that CF1 binds preferentially to GGGGTCACG by approximately a factor of 3 relative to TAGGTCACG.

Fig. 3 also shows that after only three cycles of selection, the consensus sequence was more permissive, GGGGTCAC/tG/c, where lowercase letters signify alternate base preference. A total of 71 cloned sequences showed good or reasonable matches to this consensus.

Binding to Mutated *s15* Promoters. In agreement with the above results, both CF1 and T-CF1 binding to the *s15* HRE was disrupted by clustered substitutions eliminating the upstream half-site (Fig. 4, mutation c, lane 1) but was insensitive to loss of the downstream half-site (Fig. 4, mutation b, lane 2). EMSA patterns of these factor preparations shared a major retarded band, but CF1 yielded additional lower-mobility bands that were enhanced with increasing content of long or full-length protein molecules. These bands were also insensitive to loss of the downstream half-site.

A single T → A change at position −49 corrected the imperfect symmetry of the *s15* HRE, resulting in a perfect palindrome with two inverted GGTCA repeats separated by six nucleotides (Fig. 4, PP, lane 4). This mutation did not substantially enhance the retarded EMSA bands or create additional ones, suggesting that even with this correction, the sequence context or orientation of the downstream half-site does not permit binding of additional CF1/USP molecules. In contrast, two substitutions that together convert this inverted TGTCC half-site into a direct repeat, GGTCA, enhanced the binding and led to the appearance of a novel complex of lower mobility (Fig. 4, DR6, lane 5). This apparent tandem binding of the factor was sensitive to spacing: it occurred on DR6, which has the normal 6-bp spacing between half-sites, but not on DR3, where the spacing is reduced to 3 bp (Fig. 4, lanes 5 and 6, respectively). These results are consistent with the known preference of RXRs for direct rather than inverted repeats of proper spacing (26, 27). However, tandem binding of CF1/USP did not appear to be cooperative, and the novel band was not as strong as those that correspond to binding at a single half-site, even when extensively purified CF1 was used. Moreover, tandem binding on DR6 was at least as

FIFTH PCR CYCLE SEQUENCES

	GGGGTCACG	AGAGGGTGCCA
G	GGGGTCACG	CCCCTCGTCC
	GGGGTCACG	GGCCAGCCCGC
GGC	GGGGTCACG	TCCGAGGG
	GGGGTCACG	GGGCCGCACTA
TGGC	GGGGTCACG	CGATTTG
	GGGGTCACG	CCGCCAACACC
	GGGGTCACG	CGTTAATCGCC
GG	GGGGTCACG	GCGCCGTAT
GCGACC	GGGGTCACG	GTGCC
GC	GGGGTCACG	TCCGGCCAG
AGAGAGGCC	GGGGTCACG	CC
TGCAGGTAG	GGGGTCACG	TG
	GGGGTCACG	GACCGGTGCCC
	GGGGTCACG	GCCCGCCCCAT
	GGGGTCACG	GCCGAGCTGCC
	GGGGTCACT	TCCCCGCCCCA
TGGAGCGGGC	GGGGTCAGG	G
	GGGGTCACC	ATCCACGGGCA
GGAGAGTGGTC	GGGGTCATC	
	GGGGTCAAG	ACGGGCCCCTG
	GGGGTCACA	CAGCGCCACCA
GGATGCACGA	GGGGTCACC	A
	GGGGTCAAG	GTCACGCTGCC
GT	GGGGTCACC	GCAGCTGCC
	GGGGTCACA	GCTACAGCCAT
TA	GGGGTCACC	GCCGTTTCG
TGGGCGGCA	GGGGTCATG	GG
	GGGGTCAAC	TGCGCGCCCCT
	GGGGTCATC	CCTAGCTGCCC
	GGGGTGACA	CTGGTGGCCCC
T	GGGGTCAGC	ATTACGTGCC
	GGGGTCATC	GTACCGCCTCA
	GGGGTCATT	CCATGCGACCA
	GGGGTCAAT	CACCGGCCCCC
GGGGTC	AAGGTCACG	CTCC
	GGGGTTGTA	TACCACCTCCT
A	CAGGTCGCA	AATTGTCCCC

THIRD CYCLE

	G	G	G	G	T	C	A	C/t	G/c
G	80	66	93	90	7	10	14	11	49
A	15	25	1	2	1	0	80	11	12
T	1	3	3	4	92	6	1	31	10
C	4	6	3	4	0	84	5	47	29
	68	71	71	71	71	71	71	71	68

Sample Size

FIFTH CYCLE

	G	G	G	G	T	C	A	C	G/c
G	94	95	100	100	0	3	5	5	55
A	3	5	0	0	0	0	95	11	13
T	0	0	0	0	100	3	0	16	8
C	3	0	0	0	0	94	0	68	24
	38	38	38	38	38	38	38	38	38

Sample Size

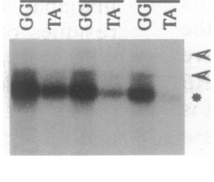

FIG. 3. Characterization of the optimal CF1/USP binding site. (*Left*) Sequences of the central 20 nucleotides of 38 independent clones containing high-affinity binding sites for CF1 after five cycles of EMSA/PCR selection on a population containing a random 20-mer flanked by defined primer sites. (*Right Upper*) Tables showing the consensus sequence of the optimal CF1 binding site, the percentile representation of each base at the indicated position, and the sample size of the independent clones used to derive these data after three or five cycles of EMSA/PCR selection. (*Right Lower*) EMSA gel. Probes were oligonucleotides representing a mutant variant of the *s15* promoter containing the optimal nonamer GGGGTCACG (lanes GG) or the wild-type nonamer TAGGTCACG (lanes TA). The relative amounts of CF1 protein used in lanes from left to right were 200 ng, 100 ng, and 50 ng. An asterisk indicates the main complex observed, and arrowheads indicate complexes of lower mobility, probably due to less degraded protein.

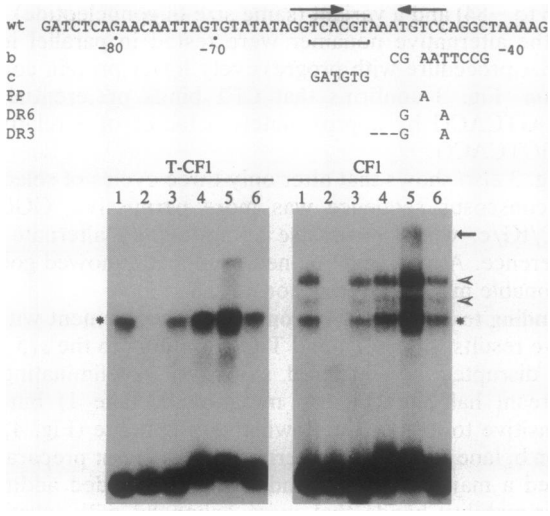

FIG. 4. Binding of CF1/USP to mutant derivatives of the *s15* regulatory region. (*Upper*) The sequence of the indicated mutants is represented under the wild-type (wt) *s15* promoter sequence. Mutant b contains clustered substitutions that destroy the right half of the palindromic site. Mutant c contains substitutions in the conserved hexamer that destroy part of the left half of the palindrome. The mutation PP has a single substitution that makes the palindrome perfectly symmetrical. Mutant DR6 contains two point mutations resulting in a direct repeat of the half-site with six bases in the spacer region, while DR3 has a direct repeat with only 3 bases in the spacer region. (*Lower*) EMSA analysis with bacterially produced CF1 (*Right*) or T-CF1 (*Left*) and probes representing the wild-type sequence (lane 3) or mutants b (lane 1), c (lane 2), PP (lane 4), DR6 (lane 5), and DR3 (lane 6). Open arrowheads indicate lower mobility complexes probably due to CF1 species that if not full length are close to it, and a dot indicates a complex due to extensively proteolyzed CF1. An arrow indicates a complex representing a putative dimeric form specific for the mutant DR6 probe.

prominent with T-CF1, despite the absence of the COOH-terminal half of the protein.

Heteromerization. CF1/USP shows substantial conservation with reference to vertebrate RXRs not only in the DNA-binding domain but also in the entire COOH-terminal half, including the putative high-affinity dimerization domain (15, 28). Since the vertebrate homologues are known to heterodimerize with THR, RAR, and vitamin D receptor (23, 28–31), we explored the possibility that the *Drosophila* protein may also heterodimerize.

We have shown that indeed CF1 can form complexes with rat THRα (rTHRα) and that this interaction cannot occur with T-CF1. When rTHRα was mixed with CF1 in the presence of the wild-type *s15* promoter, novel and strong bands of low mobility appeared (Fig. 5*A*). These bands were identified as CF1–rTHRα–DNA complexes by Western blotting of the EMSA gels: while the bands seen with CF1 alone contain CF1/USP, the novel complexes contained rTHRα as well as CF1/USP (Fig. 5, *B* and *C*). For these complexes to form with the DNA, binding of CF1 to the upstream half-site is essential, whereas DNA binding of rTHRα (to the downstream half-site as shown by footprinting analysis) is helpful but nonessential (see Fig. 5*D*, C + Tα): the complexes were abundant with the wild-type *s15* sequence (lane 1), reduced by the b mutation (lane 3), and totally absent with mutation c (lane 2). Similar lower mobility complexes were seen when the same factors were used with a fragment of the malic enzyme promoter as the DNA target (data not shown). Comparable experiments have also shown heteromerization between CF1/USP and hTHRβ or hRARα (Fig. 5*E*). The heteromerization is specific, as it is not observed between CF1/USP and hRXRα (Fig. 5*E*). These heterocomplexes form both on the *s15* HRE and on an oligonucleotide encompassing a regulatory region of the human apolipoprotein AII promoter.

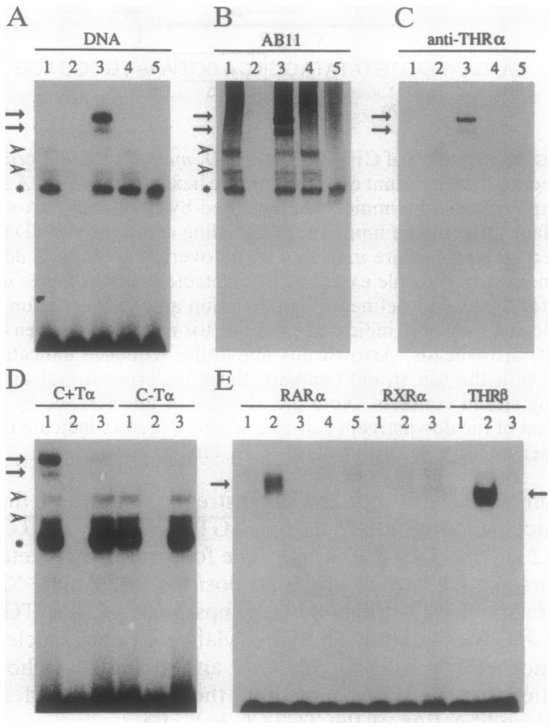

FIG. 5. Heteromerization of CF1/USP with mammalian nuclear receptors. (*A–C*) Autoradiogram of an EMSA gel with a labeled oligonucleotide corresponding to the wild-type regulatory region of the *s15* promoter as probe (*A*) and Western blots of the same gel with antibodies directed against CF1 (mAb AB11) (*B*) or against THRα (*C*). The probe was incubated with CF1 protein produced in bacteria (lane 1), extract from Sf9 cells expressing rat THRα (lane 2), CF1 plus rTHRα (lane 3), CF1 plus control Sf9 cell extract (lane 4), or T-CF1 plus rTHRα (lane 5). Arrowheads and arrows are as in Fig. 4. (*D*) EMSA gels with oligonucleotide probes corresponding to the wild-type (lane 1), mutant c (lane 2), or mutant b (lane 3) derivatives of the *s15* regulatory region. The probes were incubated with CF1 protein in the presence of extracts from Sf9 cells expressing rat THRα (lanes C+Tα) or in the presence of control Sf9 extracts (lanes C-Tα). (*E*) EMSA gels with an oligonucleotide probe corresponding to a regulatory region of the human apolipoprotein AII promoter. The probe was incubated with extracts from COS cells expressing CF1 or one of the indicated human nuclear receptor proteins. (*E Left*) CF1 extract (lane 1), CF1 plus RARα (lane 2), RARα (lane 3), control COS cell extract (lane 4), and RARα plus control COS cell extract (lane 5). (*E Center*) RXRα (lane 1), RXRα plus CF1 (lane 2), and RXRα plus control COS cell extract (lane 3). (*E Right*) THRβ (lane 1), THRβ plus CF1 (lane 2), and THRβ plus control COS cell extract (lane 3).

DISCUSSION

Ovarian Distribution of CF1/USP. It is impossible to prove formally that any antibody is absolutely specific: it could always cross-react with an as yet unknown protein. Clearly, however, the DH9 mAb is extremely specific. Its binding properties to various CF1/USP preparations establish that this antibody recognizes an epitope very near the NH_2 terminus of the protein, probably within the first 14 residues. That sequence is effectively unique. In a search of a 20-million residue nonredundant protein data base, only two marginal similarities were encountered with the first 19 residues of USP/CF1 (identity of five contiguous residues with *Aplysia* neuroactive polyproteins R15-1 and R15-2). The NH_2-terminal sequence is not shared by the RXRs of vertebrates or *B. mori* (G. Tzertzinis, personal communication), and the epitope is absent from both the *B. mori* and the *D. virilis* homologues. Thus, the immunocytochemical staining

of *D. melanogaster* follicles with DH9 (with *D. virilis* as a negative control) (Fig. 1*C*) effectively establishes that CF1/USP is present in the nuclei of follicular epithelial cells throughout oogenesis as well as in the germ-line nurse cells. A recent genetic and *in situ* hybridization study (12) documented the high abundance of usp RNA in the nurse cells but for technical reasons could not establish unambiguously whether or not the RNA is present at lower levels in follicular epithelial cells. Our finding of the factor in follicular epithelial nuclei leaves open the possibility that CF1/USP may be involved directly in transcription of the *s15* chorion gene.

DNA Binding of CF1/USP Alone. From the results presented, it is clear that bacterially produced CF1/USP can bind alone to a single HRE half-site. This property is unusual, although not unprecedented: nerve growth factor-induced protein (NGFI-B) also binds to a cognate DNA sequence bearing a single half-site (32). Although we cannot exclude the possibility that the CF1/USP binding species is dimeric, with one protein molecule making specific DNA contacts and the other not, the ready formation of complexes with T-CF1 (which lacks the probable dimerization domain) argues otherwise. Retarded bands of lower mobility form with the partially purified full-length CF1 preparation, but comparison of their mobilities with those of known dimers (THR, RAR) suggests that these bands probably are also monomeric. Whether they actually contain full-length CF1/USP or slightly degraded derivatives is as yet unknown.

To form complexes bearing additional CF1/USP molecule(s), it was necessary to convert the *s15* HRE into direct repeats spaced 6 bp apart. Even then the complexes did not appear to be cooperative; they also did not require the COOH-terminal dimerization domain.

Heteromerization. In strong contrast to the resistance of CF1/USP towards interacting with itself, we have obtained unambiguous evidence that this factor can interact cooperatively with three different mammalian hormone receptors to form complexes with target DNA. Strictly speaking the evidence is for protein heteromerization, but most likely these complexes are heterodimers. This interaction is not an artifact of the *s15* HRE, since it is observed with three different DNAs. At least in the case of *s15*, the heteromeric DNA complexes with rTHRα require binding of CF1/USP to the DNA. These complexes also show a footprint, albeit weak, on the downstream half-site; moreover, they are reduced by disrupting this half-site, suggesting some specific interaction between the DNA and the THRα partner. Cross-linking experiments indicate the formation of CF1-rTHRα heteromers even in the absence of DNA (data not shown).

The spatial requirements for heteromerization appear to be flexible: the complexes form not only on the wild-type *s15* HRE and its perfectly palindromic variant but also on DR3 (which supports neither CF1/USP nor THRα homodimers); however, DR6 appears to be optimal for heteromerization.

The observed heteromerization of fly and mammalian factors indicates that the pertinent features of RXRs have persisted through at least 500 million years of evolution, since the last common ancestor of insects and vertebrates (29). This was previously hypothesized (27) and testifies to the functional importance of this RXR property. The advantages of *Drosophila* genetics should permit detailed analysis of the functional significance of heteromeric RXR interactions, once the fly partners of CF1/USP have been identified. If CF1/USP has a function in *s15* expression it is likely to be permissive, since the protein is present in follicular nuclei long before choriogenesis. An unknown partner binding to the downstream half-site may be involved in fine-tuning chorion gene expression, since a mutation encompassing that site can disrupt the temporal specificity of *s15* transcription (M. J. Conboy, B. D. Mariani and F.C.K., unpublished results).

Note Added in Proof. Recently, CF1/USP heteromerization, notably with ecdysone receptors, has been observed by others (ref. 33; also M. R. Koelle, M. Arbeitman, and D. S. Hogness, personal communication).

We are grateful to J. Gogos and G. Tzertzinis for discussions and to M. Shea for T-CF1. We thank Drs. C. Cladaras, M. Hadzopoulou-Cladaras, and V. I. Zannis for discussions, generous supply of reagents, and advice. Additionally, we thank S. Pedreira, B. Klumpar, and T. Chai for assistance. This work has been supported by grants from the National Science Foundation and the National Institutes of Health to F.C.K. and from the National Institutes of Health to S.A.M.; A.M.K.C. was supported by a National Institutes of Health postdoctoral fellowship.

1. Segraves, W. A. (1991) *Cell* **67**, 225–228.
2. Oro, A. E., McKeown, M. & Evans, R. M. (1992) *Curr. Opin. Genet. Dev.* **2**, 269–274.
3. Ladias, J. A. A. & Karathanasis, S. K. (1991) *Science* **251**, 561–565.
4. Mangelsdorf, D. J., Umesono, K., Kliewer, S. A., Borgmeyer, U., Ong, E. S. & Evans, R. M. (1991) *Cell* **66**, 555–561.
5. Wang, L.-H., Tsai, S. Y., Cook, R. G., Beattle, W. G., Tsai, M.-J. & O'Malley, B. W. (1989) *Nature (London)* **340**, 163–166.
6. Kumar, V. & Chambon, P. (1988) *Cell* **55**, 145–156.
7. Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlman, K., Gustafsson, J.-A., Tsai, M.-J. & O'Malley, B. W. (1988) *Cell* **55**, 361–369.
8. Forman, B. M. & Samuels, H. H. (1990) *New Biol.* **2**, 587–594.
9. Shea, M. J., King, D. L., Conboy, M. J., Mariani, B. D. & Kafatos, F. C. (1990) *Genes Dev.* **4**, 1128–1140.
10. Oro, A. E., McKeown, M. & Evans, R. M. (1990) *Nature (London)* **347**, 298–301.
11. Henrich, V. C., Sliter, T. J., Lubahn, D. B., MacIntyre, A. & Gilbert, L. I. (1990) *Nucleic Acids Res.* **18**, 4143–4148.
12. Oro, A. E., McKeown, M. & Evans, R. M. (1992) *Development* **115**, 449–462.
13. Levin, A. A., Sturzenbecker, L. J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M., Lovey, A. & Grippo, J. F. (1992) *Nature (London)* **355**, 359–361.
14. Heyman, R. A., Mangelsdorf, D. J., Dyck, J. A., Stein, R. B., Eichele, G., Evans, R. M. & Thaller, C. (1992) *Cell* **68**, 397–406.
15. Mangelsdorf, D. J., Borgmeyer, U., Heyman, R. A., Zhou, J. Y., Ong, E. S., Oro, A. E., Kakizuka, A. & Evans, R. M. (1992) *Genes Dev.* **6**, 329–344.
16. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1074–1078.
17. Studier, F. W., Rosenburg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–131.
18. Smith, D. B. & Johnson, K. S. (1988) *Gene* **67**, 31–40.
19. Pollock, R. & Treisman, R. (1990) *Nucleic Acids Res.* **18**, 6197–6204.
20. Blackwell, T. K. & Weintraub, H. (1990) *Science* **250**, 1104–1110.
21. Gogos, J. A., Hsu, T., Bolton, J. & Kafatos, F. C. (1992) *Science* **257**, 1951–1955.
22. Ladias, A. A. J., Hadzopoulou-Cladaras, M., Kardassis, D., Cardot, P., Cheng, J., Zannis, V. I. & Cladaras, C. (1992) *J. Biol. Chem.* **267**, 15849–15860.
23. Hallenbeck, P. L., Marks, M. S., Lippoldt, R. E., Ozato, K. & Nikodem, V. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5572–5576.
24. Gogos, J. A., Tzertzinis, G. & Kafatos, F. C. (1991) *Nucleic Acids Res.* **19**, 1449–1453.
25. Truss, M., Chalepakis, G. & Beato, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7180–7184.
26. Nääar, A. M., Boutin, J.-M., Lipkin, S. M., Yu, V. C., Holloway, J. M., Glass, C. K. & Rosenfeld, M. G. (1991) *Cell* **65**, 1267–1279.
27. Umesono, K., Murakami, K. K., Thompson, C. C. & Evans, R. M. (1991) *Cell* **65**, 1255–1266.
28. Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J.-Y., Staub, A., Garnier, J.-M., Mader, S. & Chambon, P. (1992) *Cell* **68**, 377–395.
29. Laudet, V., Hänni, C., Coll, J., Catzeflis, F. & Stéhelin, D. (1992) *EMBO J.* **11**, 1003–1013.
30. Zhang, X.-K., Hoffmann, B., Tran, P. B.-V., Graupner, G. & Pfahl, M. (1992) *Nature (London)* **355**, 441–446.
31. Kliewer, S. A., Umesono, K., Mangelsdorf, D. J. & Evans, R. M. (1992) *Nature (London)* **355**, 446–449.
32. Wilson, T. E., Paulsen, R. E., Padgett, K. A. & Milbrandt, J. (1992) *Science* **256**, 107–110.
33. Yao, T.-P., Segraves, W. A., Oro, A. E., McKeown, M. & Evans, R. M. (1992) *Cell* **71**, 63–72.