Direct Photoaffinity Labeling of Cellular Retinoic Acid-Binding Protein I (CRABP-I) with *all-trans*-Retinoic Acid: Identification of Amino Acids in the Ligand Binding Site†

Guangping Chen and Anna Radominska-Pandya*

Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205

Received April 5, 2000; Revised Manuscript Received July 12, 2000

ABSTRACT: Cellular retinoic acid-binding proteins I and II (CRABP-I and -II, respectively) are transport proteins for *all-trans*-retinoic acid (RA), an active metabolite of vitamin A (retinol), and have been reported to be directly involved in the metabolism of RA. In this study, direct photoaffinity labeling with [11,-12-^{3}H]RA was used to identify amino acids comprising the ligand binding site of CRABP-I. Photoaffinity labeling of CRABP-I with [^{3}H]RA was light- and concentration-dependent and was protected by unlabeled RA and various retinoids, indicating that the labeling was directed to the RA-binding site. Photolabeled CRABP-I was hydrolyzed with endoproteinase Lys-C to yield radioactive peptides, which were separated by reversed-phase HPLC for analysis by Edman degradation peptide sequencing. This method identified five modified amino acids from five separate HPLC fractions: Trp7, Lys20, Arg29, Lys38, and Trp109. All five amino acids are located within one side of the "barrel" structure in the area indicated by the reported crystal structure as the ligand binding site. This is the first direct identification of specific amino acids in the RA-binding site of CRABPs by photoaffinity labeling. These results provide significant information about the ligand binding site of the CRABP-I molecule in solution.

all-trans-Retinoic acid (RA)[1] is an active metabolite of vitamin A (retinol) (*1*). It is an essential component of many biological processes such as development, cell growth, cell differentiation, and morphogenesis (*2*, *3*). RA and its synthetic analogues have also been shown to be therapeutically effective in the treatment of cancers (*4*−*7*).

The actions of RA are mediated by two classes of proteins. The first class is a family of receptor proteins that include nuclear retinoic acid receptors (RARs) and nuclear retinoid X receptors (RXRs). RARs and RXRs are retinoid-inducible transcriptional regulatory proteins that transduce the retinoid signal by altering the rate of gene transcription. The second class of proteins that mediates the actions of RA consists of small, soluble transport proteins, known as cellular retinoic acid-binding proteins (CRABP-I and CRABP-II). It has been reported that CRABP-I and CRABP-II may be directly involved in the metabolism of bound RA (*8*−*10*). A recent review (*9*) has focused on the interactions between CRABPs and the enzymes involved in retinoid metabolism, emphasizing that there is a unique and integral relationship between CRABPs and RA-metabolizing enzymes. A recent article (*11*) demonstrated both physical and functional interactions between CRABP and the RA-dependent nuclear complex (RAR).

The first primary structure of a CRABP was reported in 1985 (*12*). The first crystal structures of holo-CRABP-I and holo-CRABP-II (both complexed with RA) were determined by Kleywegt et al. (*13*). In those studies, a structure for human CRABP-II complexed with a synthetic retinoid was also reported. In addition, a recombinant form of murine apo-CRABP-I has been purified and crystallized (*14*). Both crystal structures are composed of two orthogonally packed five-stranded β-sheets and two short α-helices. The RA binding site is located deep inside the interior of the β-barrel formed by the two β-sheets. For CRABP-II, each unit cell contained one molecule. For CRABP-I, each unit cell contained two molecules, termed molecules A and B. The structures of molecules A and B in CRABP-I differ from each other. The structures of holo- and apo-CRABP-I are also different (*15*). In the unit cell of apo-CRABP-I, a β-sheet strand from molecule A forms an intermolecular β-sheet with a β-sheet of the opposing molecule B (*14*).

A crystal structure of an unliganded mutant, R111M, for human CRABP-II [apo-CRABPIIe (R111M)] has been determined (*16*). The information revealed from this apo structure, coupled with results from crystal structure-based molecular modeling studies, has resulted in a working hypothesis that postulates a three-step mechanism of RA entry into the CRABP-II molecule. This model addresses the opening of the RA entrance, the electrostatic potential that directs entry of RA into the binding pocket, and the intramolecular interactions that stabilize the RA−CRABP-

† This work was supported in part by NIH Grants DK51971, DK49715, and DK56226.

* To whom correspondence should be addressed. Telephone: (501) 686-5414. Fax: (501) 603-1146. E-mail: radominskaanna@exchange.uams.edu.

[1] Abbreviations: CRABP-I, cellular retinoic acid-binding protein I; CRBP, cellular retinol-binding protein; PDB, Protein Data Bank; RA, *all-trans*-retinoic acid; RAR, nuclear retinoic acid receptor; RXR, nuclear retinoid X receptors; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid.

10.1021/bi000321n CCC: $19.00
Published on Web 09/14/2000

II complex by locking three flexible structural elements upon RA binding to CRABP-II.

Structure−function studies of the RA binding site of CRABPs are relatively limited. Site-directed mutagenesis of conserved amino acids, such as Arg111 and Arg131 for CRABP-I and Arg111 and Arg132 for CRABP-II, has demonstrated that these residues are critical for RA binding (*17*−*19*). These results are supported by X-ray crystal structures. The two arginine residues are in contact with the carboxyl group of RA for both CRABP-I and -II. The positive charge on the functional group of arginine is important for the binding of RA. To our knowledge, structural and functional analysis of other residues has not been reported, except for information gained from the reported crystal structures and protein sequence alignments.

It has been demonstrated previously that [^{3}H]RA can be used to directly photolabel CRABP-I and -II (*20*). Recently, we utilized photoaffinity labeling to examine known ligands of CRABP-I and -II and to investigate several potential ligands. This approach resulted in the identification of two new ligands for CRABP-I and -II, 5,6-epoxy-RA and *all-trans*-retinoic acid glucuronide. The purpose of the work presented here was to use photoaffinity labeling to identify the amino acids comprising the RA binding site of CRABP-I in solution. CRABP-I was photolabeled with [11,12-^{3}H]RA and digested with proteases, and the resulting peptides were purified by HPLC and sequenced by Edman degradation. Five amino acids were identified that covalently bound RA. All five amino acids are localized to one side of the "barrel" structure which is within the area indicated by the proposed crystal structure as the ligand binding site. This is the first direct identification of specific amino acids comprising the RA binding site of CRABPs by photoaffinity labeling. The results obtained in this study will provide valuable information about the ligand binding site of the CRABP-I molecule in solution.

EXPERIMENTAL PROCEDURES

Materials. [11,12-^{3}H]-*all-trans*-Retinoic acid ([^{3}H]RA, 30 Ci/mmol) was purchased from NEN LifeScience Products (Boston, MA). Sequencing grade endoproteinase Lys-C was obtained from Roche Molecular Biochemicals (Indianapolis, IN). Tris-Bis NuPAGE gels (4 to 12%) and the NuPAGE MES SDS running buffer were purchased from Novex (San Diego, CA). All other chemicals were of the highest purity commercially available.

Preparation of Cellular Retinoic Acid-Binding Proteins. Rat cellular retinoic acid-binding protein I (CRABP-I) was provided by D. E. Ong (Vanderbilt University School of Medicine, Nashville, TN). CRABP-I was biosynthesized and purified by HPLC with an imidazole/acetate gradient (pH 6.6) as previously described (*21*). Pooled protein peaks were neutralized using Tris buffer and concentrated.

Photoaffinity Labeling of CRABP-I with [11,12-^{3}H]Retinoic Acid. CRABP-I was photolabeled with [^{3}H]RA using a method modified from that of Bernstein et al. (*20*). An aliquot of [^{3}H]RA was dried and dissolved in ethanol to a final concentration of 66 μM. CRABP-I was diluted to 0.05 mg/mL with buffer [50 mM HEPES (pH 7.5)], and 9 μL of this solution was used per assay. The reaction mixtures (10 μL final volume) were incubated for 2 min on ice, and 0.5 μL of [^{3}H]RA (3.3 μM final concentration) was added and the mixture incubated on ice for an additional 5 min. For controls (zero concentration), 0.5 μL of ethanol was added. The samples were irradiated with a long-wave (366 nm) UV light source (Spectroline model ENF-260C, Spectronics Corp., Westbury, NY) for 10 min on ice.

Following irradiation, 4 μL of 4× NuPAGE denaturing buffer (Novex, San Diego, CA) was added to each sample. Samples were vortexed, sonicated for 3 min, heated at 100 °C for 5 min, and then applied to a 1 mm NuPAGE 4 to 12% Bis-Tris gel (Novex). Following electrophoresis, the gels were stained with Coomassie Blue, destained [10% ethanol (95%)/10% glacial acetic acid], washed with water, treated with Autofluor (National Diagnostics, Manville, NJ) for 30 min, and dried. Dried gels were subjected to autoradiography at −80 °C for 2−5 days. Autoradiographs were analyzed and quantified by densitometry using an IS-1000 Digital Imaging System (version 2.03, Alpha Innotech Corp.).

Proteolytic Hydrolysis of [^{3}H]RA-Labeled CRABP-I with Endoproteinase Lys-C. To evaluate the proteolytic hydrolysis of CRABP-I after photoaffinity labeling with [^{3}H]RA, 15 μg of CRABP-I in 0.3 mL of 50 mM HEPES was labeled with 3.3 μM [^{3}H]RA as described above. The protein was precipitated with TCA at a final concentration of 10%, and samples were centrifuged at 15000*g*. The pellets were washed once with ice-cold ethanol and then dissolved in 30 μL of buffer [50 mM Tris (pH 8.0) and 0.1% SDS]. Endoproteinase Lys-C (0.5 μg) was added, and the mixture was incubated at room temperature. Aliquots (5 μL) were removed from the mixture at 0, 15, 30, and 60 min and 15 h and added to 20 μL of SDS−PAGE denaturing buffer. Labeled peptides in these aliquots were separated on a 4 to 12% Bis-Tris NuPAGE gel and were visualized by autoradiography as described above.

To generate radioactive peptides for Edman degradation, 100 μg of CRABP-I in 0.5 mL of 50 mM HEPES (pH 7.4) (13 μM) was photolabeled with 1.3 μM [^{3}H]RA for 10 min on ice as described above. Cold RA (30 μM) was then added to the mixture, and the sample was irradiated for an additional 10 min on ice. Following the labeling, 500 μL of a charcoal/dextran solution [1% activated charcoal and 0.1% dextran (w/v) in 10 mM Tris, 0.25 M sucrose, and 1 mM EDTA (pH 7.4)] was added. After shaking at 4 °C for 10 min, the charcoal was separated by centrifugation. The supernatant, containing the protein, was precipitated with a final concentration of 10% TCA. The protein that precipitated was collected by centrifugation, washed with cold ethanol, and dissolved in 100 μL of buffer [50 mM Tris (pH 8.0) and 0.1% SDS]. Three micrograms of sequencing grade endoproteinase Lys-C was added, and the mixture was incubated at room temperature for 48 h before HPLC separation was carried out. A control hydrolysis including everything except CRABP-I was carried out under the exact same conditions.

HPLC Separation of Peptides from Endoproteinase Lys-C-Hydrolyzed Photolabeled CRABP-I. For HPLC separation, a reversed-phase C18 column was used. Gradient solvent A contained 2% HPLC grade acetonitrile and 0.06% trifluoroacetic acid in HPLC grade water; solvent B contained 70% HPLC grade acetonitrile and 0.06% trifluoroacetic acid in HPLC grade water. A linear gradient from 98% A to 98% B was used from 0 to 100 min at a flow rate of 1 mL/min. The column elute was monitored for UV absorption at 215 nm. A control to establish a background for the HPLC

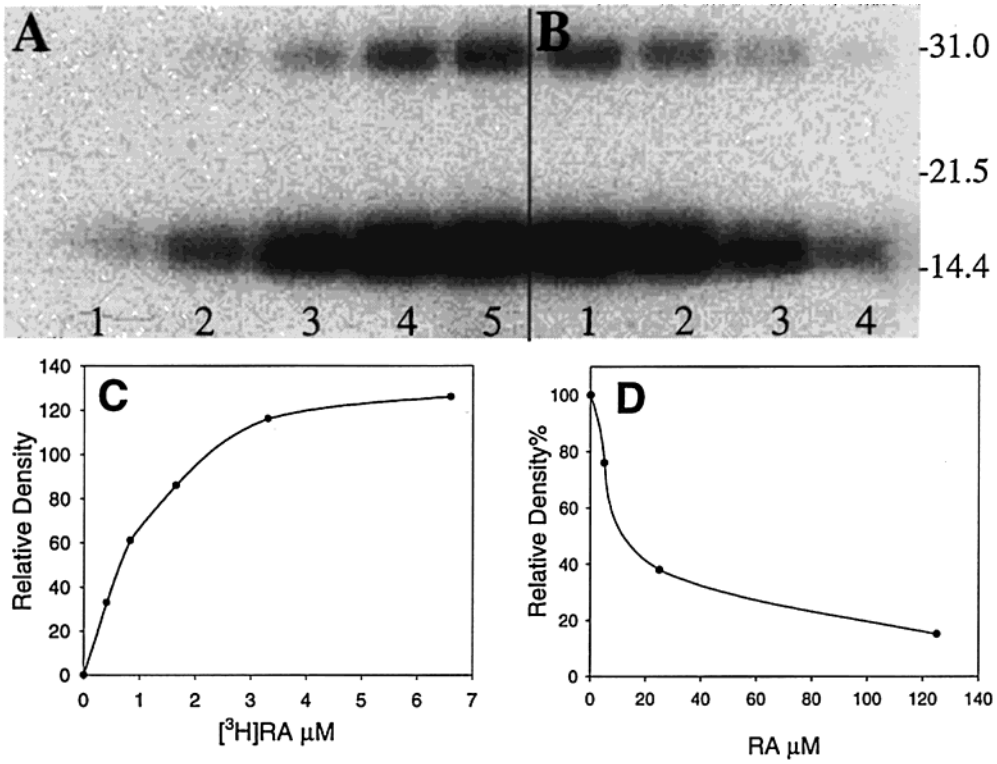


FIGURE 1: Photoaffinity labeling of CRABP-I by *all-trans*-retinoic acid. (A) Concentration-dependent photoaffinity labeling of CRABP-I. CRABP-I (3.1 μM) in 50 mM HEPES (pH 7.4) was incubated with (1) 0.42, (2) 0.83, (3) 1.6, (4) 3.3, and (5) 6.6 μM [^{3}H]RA on ice followed by exposure to UV 366 nm light for 10 min on ice. The samples were subjected to SDS−PAGE, and the radioactive bands on the autoradiograph of the gel are shown. (B) Unlabeled RA protection of CRABP-I from photoaffinity labeling by [^{3}H]RA. CRABP-I (3.1 μM) in 50 mM HEPES (pH 7.4) was incubated with different concentrations of cold RA on ice: (1) 0, (2) 5, (3) 25, and (4) 125 μM. The mixtures were then photolabeled with 3.3 μM [^{3}H]RA. (C) Quantitative relationship between relative density and [^{3}H]RA concentration. The calculated half-saturation was 1.7 μM. (D) Quantitative relationship for the protection by unlabeled RA of the photoaffinity labeling of CRABP-I with [^{3}H]RA. The densitometry was analyzed using an IS-1000 Digital Imaging System (version 2.03, Alpha Innotech Corp.).

profile was run using the control hydrolysis. Fractions of 1 mL were collected, and an aliquot of 0.1 mL from each fraction was used for scintillation counting.

Edman Degradation of Covalently Modified Peptides. The fractions from the radioactive peaks of the HPLC profile were subjected to automated N-terminal protein sequencing using an Applied Biosystems 477A protein sequencer at the Protein Microanalysis Facility at The University of Texas at Austin.

RESULTS

Characterization of the Direct Photoaffinity Labeling of CRABP-I by [11,12-^{3}H]RA. [^{3}H]RA has been used as a direct photoaffinity labeling probe for CRABPs (*20*) and microsomal UDP-glucuronosyltransferases (*22*). Figure 1 demonstrates that CRABP-I was directly photolabeled by [^{3}H]RA. The labeling was concentration-dependent (Figure 1A) and protected by unlabeled RA (Figure 1B). The semiquantitative graph of density versus [^{3}H]RA concentration demonstrated that RA binds to CRABP-I at a ratio of 1/1. When 3.1 μM CRABP-I was used, half-saturation was reached at 1.7 μM [^{3}H]RA. The labeling was also UV light (366 nm)-dependent and protected by other retinoids (data not shown). CRABP-I photolabeled with [^{3}H]RA was subjected to hydrolysis with endoproteinase Lys-C. Peptides were separated by SDS−PAGE with some of the peptides retaining the radioactive label (Figure 2), providing further proof that [^{3}H]RA was covalently linked to amino acid residues of CRABP-I.

HPLC Separation of the Radioactive Peptides Derived from Endoproteinase Lys-C-Catalyzed Hydrolysis of [^{3}H]RA-Labeled CRABP-I. CRABP-I photolabeled with [^{3}H]RA was hydrolyzed with endoproteinase Lys-C for 48 h. Following hydrolysis, peptides were separated by reversed-phase HPLC and 1 mL fractions were collected. An aliquot of 0.1 mL from each fraction was used for scintillation counting (radioactive profile for the HPLC separation shown in Figure 3B). Because of the high molecular weight and hydrophobicity of the RA ligand, the labeled peptides eluted from the column later in the gradient (higher acetonitrile concentration) than did the unlabeled peptides. Peptides that eluted from the column before 60 min did not contain radioactive peptides.

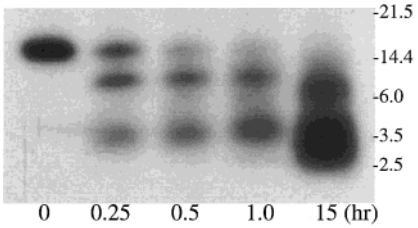


FIGURE 2: Separation of peptides from endoproteinase Lys-C-hydrolyzed photolabeled CRABP-I by gel electrophoresis. CRABP-I (15 μg) in 0.3 mL of 50 mM HEPES (pH 7.0) was labeled with 3.3 μM [^{3}H]RA. The labeled CRABP-I was hydrolyzed with endoproteinase Lys-C. Two microgram samples of CRABP-I were taken at 0, 15, 30, and 60 min and 15 h and subjected to electrophoresis on a 4 to 12% Bis-Tris NuPAGE gel. The autoradiograph of the gel is shown.

Peptide Sequencing. Fractions 61, 64, 66, 69, and 72 of the radioactivity profile (Figure 3B) were selected for peptide

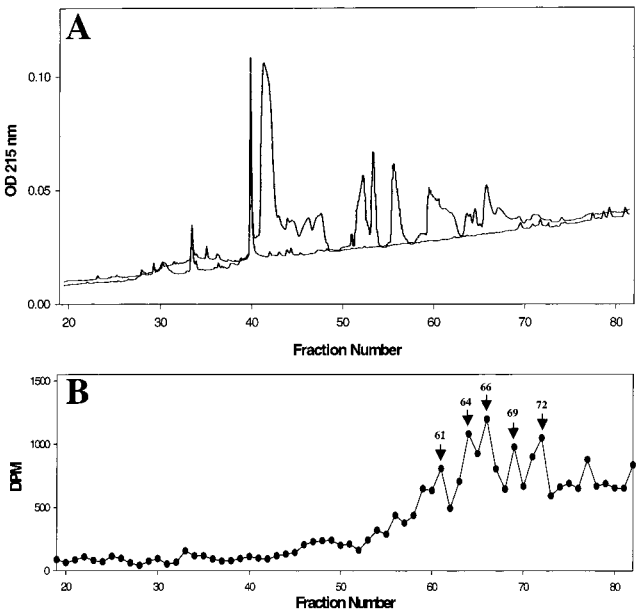


FIGURE 3: HPLC profile for the separation of radioactive peptides. [^{3}H]RA-labeled CRABP-I (100 μg) was hydrolyzed with sequencing grade endoproteinase Lys-C. The hydrolyzed peptides were separated on a reversed-phase C18 column. The column eluent was monitored by UV absorption at 215 nm (A). A control (hydrolysis without CRABP-I) with few peaks is also shown in panel A. Fractions of 1 mL (flow rate, 1 mL/min) were collected, and an aliquot of 0.1 mL from each fraction was used for scintillation counting (B). The fractions at the maximum of the radioactivity peaks in the HPLC profile (fractions 61, 64, 66, 69, and 72) were used for N-terminal protein sequencing.

sequencing. The N-terminal protein sequencing was carried out at the Protein Microanalysis Facility at The University of Texas at Austin. Sequencing results from these fractions are summarized in Table 1. Fraction 61 contained one polypeptide, consisting of 20 amino acids that correspond to residues 21−40 of the CRABP-I primary sequence. The amino acid corresponding to residue K38 on CRABP-I was undetectable, suggesting K38 was covalently modified by retinoic acid. Fraction 64 contained two polypeptides, corresponding to amino acids 9−20 and 21−38 of CRABP-I. The peptide of residues 9−20 was not labeled. Residue R29 in the peptide corresponding to residues 21−38 was undetectable, suggesting R29 was modified by RA. Fraction 66 also contained two polypeptides. One of them, corresponding to residues 21−38, was not labeled. In the other peptide (93−112), Residue W109 was labeled. For fraction 69, three polypeptides were found. Only W7 in the peptide corresponding to amino acids 1−8 of CRABP-I was labeled. According to the sequencing results, there were five residues labeled: W7, K20, R29, K38, and W109. The amino acid sequence of CRABP-I is shown in Table 2. The labeled residues are marked using bold letters.

Sequence Alignment. The primary sequence for rat CRABP-I was aligned with two CRABP-II sequences and two cellular retinol-binding protein (CRBP) sequences (Figure 4) by the computer program, MultAlin (*23*). The protein sequences of different CRABPs are highly conserved. CRABP-Is from mouse, bovine, and rat have identical amino acid sequences and share all but one amino acid with CRABP-I from human. The degree of amino acid conservation between proteins from the CRABP-I and CRABP-II families is also high (>70%). There is also some correlation between amino acid sequences from CRABPs and CRBPs. The degree of amino acid identity among the five proteins shown in Figure 4 is 34.6%, and the degree of similarity among the five proteins is 81.6%. Of the five amino acid residues that were labeled, four of them (W7, R29, K38, and W109) are conserved among the CRABPs and CRBPs, while the other residue (K20) is conserved in CRABPs as K and in CRBPs as R (all the amino acid numbers correspond to the CRABP-I sequence). K and R are the only two residues that carry positive charges on the side chain.

The two shaded amino acids in Figure 4 (R111 and R131) are two arginines that previously had been identified to be essential for binding RA to CRABP by both site-directed mutagenesis and the crystal structure. The two residues are conserved in CRABPs as R, but are conserved in CRBPs as Q. The positive charge on R determines the specificity of CRABPs for binding the carboxyl group of retinoic acid.

DISCUSSION

Although several crystal structures for CRABP-I and -II have been determined, there are a limited number of studies in which the relationship between CRABP structure and function is examined (*15*, *18*−*20*, *24*−*26*). Crystal structure studies on CRABP-II (*16*) have proposed a model for the entry of RA into the RA-binding site of CRABP-II. Even though this approach has identified potential amino acids that might be involved in RA binding, the information obtained about a protein structure in its crystal state needs to be confirmed through functional studies. Site-directed mutagenesis and direct photoaffinity labeling are two valuable techniques that can empirically confirm findings from structural studies. While there have been limited attempts to characterize the amino acids involved in RA binding to CRABP-I by site-directed mutagenesis, there have been no previous studies in which photoaffinity labeling was applied to determine which amino acids are involved in RA binding to CRABP-I.

In this study, [11,12-^{3}H]RA was utilized to identify specific amino acids localized in the RA-binding site of recombinant CRABP-I. We focused our attention on identifying specific amino acids involved in RA binding and comparing these findings with information available from the reported CRABP-I X-ray crystal coordinates. By using photoaffinity labeling, we are able to provide valuable information that complements the structural data provided by X-ray crystal structures. Our photoaffinity labeling method provides direct evidence of amino acid residues that are important for ligand binding in actual solution. This is an inexpensive but powerful method for identification of crucial active site amino acids. Moreover, it can also be used for experimental testing of new ligands and inhibitors. We have also used this method as a first step in characterizing mutants obtained by site-directed mutagenesis.

For comparison of the amino acids identified here by photoaffinity labeling with the crystal structure, we selected the bovine holo-CRABP-I crystal structure complexed with RA. This crystal structure was determined in 1994 at a resolution of 2.9 Å (PDB entry 1CBR) (*27*). The bovine CRABP-I amino acid sequence is identical to those of rat and mouse CRABP-I, and it is only one amino acid different from human CRABP-I. Figure 5 shows the crystal structure using RasMol 2.5 program. The basic crystal structure in Figure 5 is displayed as "ribbons", and the ligand retinoic

Table 1: Amino Acid Sequences Obtained by Edman Degradation[a]

#[b]	aa#[c]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
61	21-40	A	L	G	V	N	A	M	L	R	K	V	A	V	A	A	A	S	-	P	H
	pmol	42	23	13	26	1.6	16	11	8.1	0.7	1.9	10	7.6	9.7	7.4	8.5	8.5	0.4	0	1.5	0.1
64	09-20	M	R	S	S	E	N	F	D	E	L	L	K								
	pmol	15	63	4.5	5.4	3.4	2.2	1.7	1.0	5.9	2.3	0.9	0.4								
	21-38	A	L	G	V	N	A	M	L	-	K	V	A	V	A	A	A	S	K		
	pmol	3.0	2.3	1.6	1.4	1.3	1.2	0.8	3.0	0	0.6	0.9	0.4	1.1	2.7	3.4	3.1	0.9	0.4		
66	93-112	I	H	(C)	T	Q	T	L	L	E	G	D	G	P	K	T	Y	-	T	R	E
	pmol	2.9	3.7	?	2.9	1.7	3.8	2.0	4.0	3.7	1.2	1.3	2.4	1.1	0.5	0.8	0.7	0	1.1	1.7	0.7
	21-38	A	L	G	V	N	A	M	L	R	K	V	A	V	A	A	A	S	K		
	pmol	2.4	1.3	1.7	1.1	1.0	1.8	0.8	3.9	3.2	0.3	1.1	1.2	1.1	0.9	1.3	1.4	0.3	0.2		
69	93-106	I	H	(C)	T	Q	T	L	L	E	G	D	G	P	K						
	pmol	1.2	0.7	?	1.0	0.8	0.5	0.4	0.6	0.9	1.2	0.6	0.8	0.1	0.1						
	01-08	P	N	F	A	G	T	-	K												
	pmol	0.3	0.8	0.8	0.8	1.2	0.5	0	0.1												
	21-30	A	L	G	V	N	A	M	L	R	K										
	pmol	1.1	0.9	1.0	0.4	0.3	0.5	0.2	0.6	0.6	0.2										
72	09-28	M	R	S	S	E	N	F	D	E	L	L	-	A	L	G	V	N	A	M	L
	pmol	0.8	1.6	0.1	0.1	0.5	0.2	0.2	1.4	0.8	1.3	1.3	0	0.5	0.8	1.1	0.2	0.3	0.3	0.2	0.5

[a] Note that the − sign means that no amino acid can be sequenced in that cycle. Cysteine (C) cannot be sequenced by this method. [b] Fraction number corresponding to Figure 3B. [c] Corresponding amino acid number in the CRABP-I sequence.

Table 2: Sequence of CRABP-I and the Labeled Amino Acids[a]

```
  1 PNFAGTWKMR SSENFDELLK ALGVNAMLRK VAVAAASKPH VEIRQDGDQF  50
 51 YIKTSTTVRT TEINFKVGEG FEEETVDGRK CRSLPTWENE NKIHCTQTLL 100
101 EGDGPKTYWT RELANDELIL TFGADDVVCT RIYVRE 136
```

[a] Labeled residues are bold, and amino acids that are less than 5 Å from the ligand are shaded.

```
                  1                                                  50
CRABP-I (rat)     .P.NFAGTWK MRSSENFDEL LKALGVNAML RKVAVAAASK PHVEIRQDGD
CRABP-II(hum)     MP.NFSGNWK IIRSENFEEL LKVLGVNVML RKIAVAAASK PAVEIKQEGD
CRABP-II(rat)     MP.NFSGNWK IIRSENFEEM LKALGVNMMM RKIAVAAASK PAVEIKQEND
CRBP(rat)         MPVDFNGYWK MLSNENFEEY LRALDVNVAL RKIANLL..K PDKEIVQDGD
CRBP (hum)        MPVDFTGYWK MLVNENFEEY LRALDVNVAL RKIANLL..K PDKEIVQDGD

                  51                                                100
CRABP-I (rat)     .QFYIKTSTT VRTTEINFKV GGEFEEET.. VDGRKCRSLP TWENENKIHC
CRABP-II(hum)     .TFYIKTSTT VRTTEINFKV GEEFEEQT.. VDGRPCKSLV KWESENKMVC
CRABP-II (rat)    DTFYIKTSTT VRTTEINFKI GEEFEEQT.. VDGRPCKSLV KWESENKMVC
CRBP(rat)         .HMIIRTLST FRNYIMDFQV GKEFEEDLTG IDDRKCMTTV SWDGD.KLQC
CRBP (hum)        .HMIIRTLST FRNYIMDFQV GKEFEEDLTG IDDRKCMTTV SWDGD.KLQC

                  101                                          143
CRABP-I (rat)     TQTLLEGDGP KTYWTRELAN D.ELILTFGA DDVVCTRIYV RE.
CRABP-II(hum)     EQKLLKGEGP KTSWTRELTN DGELILTMTA DDVVCTRVYV RE.
CRABP-II (rat)    EQRLLKGEGP KTSWSRELTN DGELILTMTA DDVVCTRVYV RE.
CRBP(rat)         VQ...KGEKE GRGWTQWIEG D.ELHLEMRA EGVTCKQVFK KVH
CRBP (hum)        VQ...KGEKE GRGWTQWIEG D.ELHLEMRV EGVVCKQVFK KVQ
```

FIGURE 4: Sequence alignment of CRABPs and CRBPs. The computer program MultAlin (*23*) was used for sequence alignment. The five boxed residues are labeled with [^{3}H]RA. The shaded Arg residues are conserved in CRABPs but not in CRBPs, and mutagenesis of these residues has demonstrated that they are essential for RA binding specificity. All the protein sequences are from the GenBank. The references describing each of the DNAs encoding these proteins are as follows: CRABP-I (rat), accession number M17253 (*29*, *30*); CRABP−II (hum), accession number M68867 (*31*); CRABP−II (rat), accession number U23407 (*32*); CRBP (rat), accession number M16459 (*33*); and CRBP (hum), accession number X07437 (*34*).

acid is displayed as a "space-fill" diagram. The unit structure is formed by two molecules of CRABP-I. Each of the barrel-formed structures is composed of 10 β-sheets and two α-helices. The ligand binds in the barrel near the side where the two α-helices are situated. The five labeled residues are all situated near the ligand. The nearest distances between the ligand and the five labeled residues are shown in Table 3 and are all less than 8.4 Å. This is a reasonable distance for chemical reaction to take place. Lys20 and Arg29 are located within the two α-helices and close to the ring structure of retinoic acid. Lys38 is located between an α-helix and a β-sheet, which is close to the carboxyl end of retinoic acid. Trp109 is in the same β-sheet as Arg111. Trp7 is in a β-sheet located next to the β-sheet where Arg131 is located. Both Trp109 and Trp7 are also close to the carboxyl end of retinoic acid. The five labeled amino acids are all located near the two ends of the five-double bond conjugate system (Figure 6). This agrees with the chemical properties of the

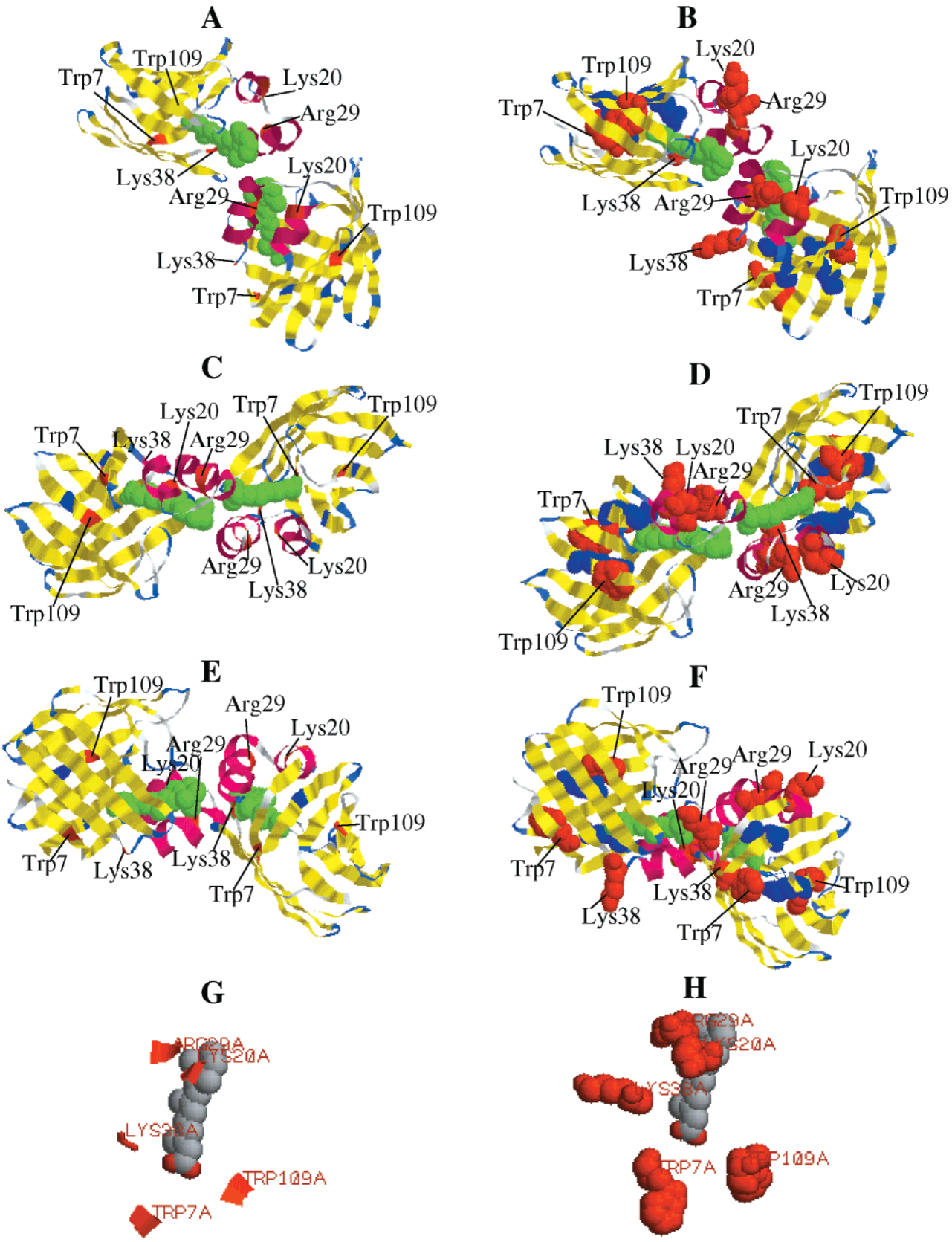
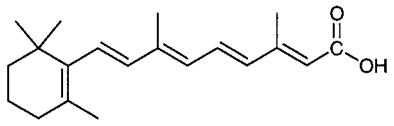


FIGURE 5: CRABP-I crystal structure showing the positions of the ligand and the residues labeled with [^{3}H]RA. CRABP-I coordinates were downloaded from the Protein Data Bank (entry 1CBR). The crystal structures were developed using the RasMol 2.5 program. The basic crystal structure is displayed as ribbons, while the ligand, retinoic acid, is displayed as a space-fill diagram. Panels A, C, and E show the structure from three different angles. Panels B, D, and F show the structure from the same three different angles as panels A, C, and E, except that the five labeled residues and Arg111 and Arg131 are shown as space-fill diagrams. Panels G and H display the ligand and the five labeled amino acids only. In panel G, the ligand was displayed as a space-fill diagram, while the five labeled amino acids are displayed as ribbons. In panel H, the ligand and five amino acids are displayed as a space-fill diagram.

Table 3: Distance between Labeled Residues and Retinoic Acid in the Crystal Structure[a]

	Trp7	Lys20	Arg29	Lys38	Trp109
shortest distance (Å)	8.1	8.3	5.9	6.4	8.4

[a] The distance is calculated according to the coordinates of CRABP-I (PDB entry 1CBR).



all-*trans*-Retinoic Acid

FIGURE 6: Chemical structure of *all-trans*-retinoic acid.

conjugate system since, when exposed to UV light, free radicals are more easily produced near the two ends of the double-bond conjugate system.

Although the five labeled residues were not listed as direct contact (with ligand) residues in the crystal structure (distance of less than 5 Å) (PDB entry 1CBR), the calculated distance (5.9−8.4 Å) proved that they are all close to the ligand (Table 3). The molecular structure of CRABP-I may be slightly different in solution and in crystal form because of the different environmental conditions and crystal packing. Crystal structures of the same CRABP-I reported from different laboratories have slight variations (*14*). The binding of RA in the CRABP-I binding pocket may also have a certain flexibility (*16*). It has been suggested that the movement of the loop region may facilitate the binding of RA in the binding pocket (*14*).

Trp7 and Trp109 have not been shown by any previous studies to be important for RA binding, but they were labeled by RA in this work. According to the crystal structure, these two residues appear to be like a door on the bottom of the barrel structure to prevent RA from going through the open bottom. The formation of intermolecular β-sheets (*16*) between CRABP-I and other proteins may slightly change the barrel structure of CRABP-I. This change may open the door and change the binding affinity of the binding site and facilitate the movement of RA from CRABP-I to other proteins. It had been reported that CRABP-I can be phosphorylated (*28*). This phosphorylation may provide the energy needed for the conformation change for the channeling of RA in vivo.

All the five labeled residues are conserved among the sequence alignment of three CRABPs and two CRBPs (Figure 4). This suggests that these five residues are important for the binding of the basic structure of retinoids. R111, R131, and other residues in CRABP-I may determine the specificity of CRABP for retinoic acid (carboxyl group) as opposed to retinol or retinal.

ACKNOWLEDGMENT

We thank Professor David E. Ong for the generous gift of purified CRABP-I. The important critique of the manuscript from Dr. John Lyndal York is deeply appreciated. The N-terminal protein sequencing was done by Dr. Klaus D. Linse at the Protein Microanalysis Facility at The University of Texas at Austin. Technical assistance for HPLC operation by Ms. Joyce S. Massengill is appreciated. We also thank Ms. Joanna Little for her critical evaluation of the manuscript.

REFERENCES

1. Napoli, J. L. (1996) *FASEB J. 10*, 993−1001.
2. Hoffman, C., and Eichelle, G. (1994) *Retinoids in Development*, 2nd ed., Raven Press, New York.
3. Gudas, L., Sporn, M., and Roberts, A. (1994) *Cellular Biology and Biochemistry of the Retinoids*, 2nd ed., Raven Press, New York.
4. Berman, E. (1997) *Curr. Opin. Hematol. 4*, 256−260.
5. Lazzarino, M., Regazzi, M. B., and Corso, A. (1996) *Leukemia Lymphoma 23*, 539−543.
6. Miller, W. H., Jr. (1998) *Cancer 83*, 1471−1482.
7. Sabichi, A. L., Lerner, S. P., Grossman, H. B., and Lippman, S. M. (1998) *Curr. Opin. Oncol. 10*, 479−484.
8. Boylan, J., and Gudas, L. (1992) *J. Biol. Chem. 267*, 21486−21491.
9. Napoli, J. (1999) *Biochim. Biophys. Acta 1440*, 139−162.
10. Zheng, W. L., Bucco, R. A., Sierra-Rievera, E., Osteen, K. G., Melner, M. H., and Ong, D. E. (1999) *Biol. Reprod. 60*, 110−114.
11. Delva, L., Bastie, J.-N., Rochette-Egly, C., Kraiba, R., Balitrand, N., Despouy, G., Chambon, P., and Chomienne, C. (1999) *Mol. Cell. Biol. 19*, 7158−7167.
12. Sundelin, J., Das, S. R., Eriksson, U., Rask, L., and Peterson, P. A. (1985) *J. Biol. Chem. 260*, 6494−6499.
13. Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K., and Jones, T. A. (1994) *Structure 2*, 1241−1258.
14. Thompson, J. R., Bratt, J. M., and Banaszak, L. J. (1995) *J. Mol. Biol. 252*, 433−446.
15. Li, E., and Norris, A. W. (1996) *Annu. Rev. Nutr. 16*, 205−234.
16. Chen, X., Tordova, M., Gilliland, G. L., Wang, L., Li, Y., Yan, H., and Ji, X. (1998) *J. Mol. Biol. 278*, 641−653.
17. Chen, L. X., Zhang, Z. P., Scafonas, A., Cavalli, R. C., Gabriel, J. L., Soprano, K. J., and Soprano, D. R. (1995) *J. Biol. Chem. 270*, 4518−4525.
18. Wang, L., Li, Y., and Yan, H. (1997) *J. Biol. Chem. 272*, 1541−1547.
19. Zhang, J., Liu, Z. P., Jones, T. A., Gierasch, L. M., and Sambrook, J. F. (1992) *Proteins 13*, 87−99.
20. Bernstein, P. S., Choi, S. Y., Ho, Y. C., and Rando, R. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 654−658.
21. Jamison, R. S., Newcomer, M. E., and Ong, D. E. (1994) *Biochemistry 33*, 2873−2879.
22. Little, J. M., and Radominska, A. (1997) *Biochem. Biophys. Res. Commun. 230*, 497−500.
23. Corpet, F. (1988) *Nucleic Acids Res. 16*, 10881−10890.
24. Martin, C. A., Dawson, M. I., McCormick, A. M., and Napoli, J. L. (1986) *Biochem. Biophys. Res. Commun. 135*, 124−130.
25. Sani, B. P., Wille, J. J., Jr., Dawson, M. I., Hobbs, P. D., Bupp, J., Rhee, S., Chao, W. R., Dorsky, A., and Morimoto, H. (1990) *Chem.-Biol. Interact. 75*, 293−304.
26. Cheng, L., Qian, S. J., Rothschild, C., d'Avignon, A., Lefkowith, J. B., Gordon, J. I., and Li, E. (1991) *J. Biol. Chem. 266*, 24404−24412.
27. Kleywegt, G., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K., and Jones, T. (1994) *Structure 2*, 1241−1258.
28. Cope, F. O., Staller, J. M., Mahsem, R. A., and Boutwell, R. K. (1984) *Biochem. Biophys. Res. Commun. 120*, 593−601.
29. Shubeita, H. E., Sambrook, J. F., and McCormick, A. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5645−5649.
30. Rajan, N., Kidd, G. L., Talmage, D. A., Blaner, W. S., Suhara, A., and Goodman, D. S. (1991) *J. Lipid Res. 32*, 1195−1204.
31. Astrom, A., Tavakkol, A., Pettersson, U., Cromie, M., Elder, J., and Voorhees, J. (1991) *J. Biol. Chem. 266*, 17662−17666.
32. Bucco, R. A., Melner, M. H., Gordon, D. S., Leers-Sucheta, S., and Ong, D. E. (1995) *Endocrinology 136*, 2730−2740.
33. Sherman, D., Lloyd, R., and Chytil, F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3209−3213.
34. Nilsson, M., Spurr, N., Lundvall, J., Rask, L., and Peterson, P. (1988) *Eur. J. Biochem. 173*, 35−44.

BI000321N