J. Mol. Biol. (1998) 278, 641–653

JMB

Crystal Structure of Apo-cellular Retinoic Acid-binding Protein Type II (R111M) Suggests a Mechanism of Ligand Entry

Xin Chen[1], Maria Tordova[2], Gary L. Gilliland[2], Lincong Wang[3], Yue Li[3] Honggao Yan[3] and Xinhua Ji[1]*

[1]*ABL-Basic Research Program National Cancer Institute-Frederick Cancer Research and Development Center, P.O. Box B, Frederick, MD 21702 USA*

[2]*Center for Advanced Research in Biotechnology of the University of Maryland Biotechnology Institute & National Institute of Standards and Technology, 9600 Gudelsky Dr., Rockville, MD 20850, USA*

[3]*Department of Biochemistry Michigan State University East Lansing, MI 48824 USA*

**Corresponding author*

The crystal structure of unliganded mutant R111M of human cellular retinoic acid-binding protein type II (apo-CRABPII (R111M)) has been determined at 2.3 Å and refined to a crystallographic *R*-factor of 0.18. Although the mutant protein has lower affinity for all-*trans*-retinoic acid (RA) than the wild-type, it is properly folded, and its conformation is very similar to the wild-type. apo-CRABPII (R111M) crystallizes in space group *P*1 with two molecules in the unit cell. The two molecules have high structural similarity except that their α2 helices differ strikingly. Analyses of the molecular conformation and crystal packing environment suggest that one of the two molecules assumes a conformation compatible with RA entry. Three structural elements encompassing the opening of the binding pocket exhibit large conformational changes, when compared with holo-CRABPII, which include the α2 helix and the βC–βD and βE–βF hairpin loops. The α2 helix is unwound at its N terminus, which appears to be essential for the opening of the RA-binding pocket. Three arginine side-chains (29, 59, and 132) are found with their guanidino groups pointing into the RA-binding pocket. A three-step mechanism of RA entry has been proposed, addressing 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·CRABPII complex *via* locking the three flexible structural elements when RA is bound.

Keywords: cellular retinoic acid-binding protein type II; retinoic acid; crystal structure; ligand entry; intracellular lipid-binding proteins

Introduction

The biological effects of retinoic acid (RA) extend far beyond those of its precursors, retinol (vitamin A) and retinal, commonly known to be critical for vision. RA has been suggested to normalize cell differentiation and growth (Elias & Williams, 1981; Gudas *et al.*, 1994; Lotan, 1980), and has been proven effective in inhibiting numerous skin disorders and human cancers (Hong & Itri, 1994; Tallman, 1996). The regulatory effects of RA are believed to be mediated through nuclear receptors (Giguere *et al.*, 1987; Mangelsdorf *et al.*, 1990; Petkovich *et al.*, 1987). The intracellular existence of RA is modulated by two kinds of RA-binding proteins, cellular RA-binding protein type I (Ong & Chytil, 1975; Sani & Hill, 1974) and type II (Bailey & Siu, 1988) (CRABPI and II). The sequences of CRABPs are highly conserved in vertebrates. For instance, CRABPIs from rat, mouse, and cow are identical, and differ by only one residue from the human sequence. CRABPI and II are closely related, with 76% identity in human. CRABPs play important roles in mediating the biological functions of RA. Instead of being solely passive carriers of RA, CRABPs have been shown to facilitate RA modification upon interaction with other proteins (Boylan & Gudas, 1992; Fiorella & Napoli, 1991, 1994;

Abbreviations used: CRABPI, cellular retinoic acid-binding protein type I; CRABPII, cellular retinoic acid-binding protein type II; CRBPII, cellular retinol-binding protein type II; I-FABP, intestinal fatty acid-binding protein; ALBP, adipocyte lipid-binding protein; iLBP, intracellular lipid-binding protein; RA, retinoic acid; rms, root mean square; vdw, van der Waals; PDB, Protein Data Bank; PEG, polyethylene glycol.

0022–2836/98/180641–13 $25.00/0/mb981734

Napoli *et al*., 1995). Recently, the unique function of CRABPs was proposed as maintaining (instead of limiting) cellular RA availability when dietary RA is in short supply, since the pivotal involvement of CRABPs in the RA signal pathway is ruled out by the fact that mice deficient in CRABP genes are essentially normal in their development (Gorry *et al*., 1994; Lampron *et al*., 1995). More recently, it was suggested that CRABPII may participate in RA production (Bucco *et al*., 1997), and that CRABPII is a positive regulator of RA signaling in breast cancer cells (Jing *et al*., 1997).

CRABPs belong to a family of proteins named intracellular lipid-binding proteins (iLBP), including fatty acid-binding proteins, lipid-binding proteins, and cellular retinoid-binding proteins. More than 25 crystal structures of these proteins are available from the Protein Data Bank (PDB). The structures have been the subject of several reviews (Banaszak *et al*., 1994; Newcomer, 1995; Ong *et al*., 1994). These proteins share a common structural motif, a β-barrel consisting of ten up-and-down antiparallel β-strands A through J (Figure 1). One end of the β-barrel is blocked by the N terminus of the protein, whereas the other end is covered by a helix-loop-helix motif, α1-loop-α2. The ligand-binding site is deep inside the β-barrel. The end of the elongated ligand at the portal surrounded by the α2 helix and two neighboring β-hairpin loops is usually solvent-accessible (Figure 1). Irrespective of the type of the ligand, the ligand-binding modes of iLBPs share a common feature in that the ligand is unable to enter or exit the pocket without eliciting significant conformational changes in the protein. The ligand-entry problem has been discussed in numerous papers (Banaszak *et al*., 1994; Winter *et al*., 1993; Xu *et al*., 1993). The common approach to this problem is to compare the apo- and holo-structures. However, it is hindered by two obstacles. First, so far both apo- and holo-structures have been determined for only four proteins: CRABPI (Kleywegt *et al*., 1994; Thompson *et al*., 1995), cellular retinol-binding protein type II (CRBPII) (Winter *et al*., 1993), intestinal fatty acid-binding protein (I-FABP) (Sacchettini *et al*., 1989a,b; Scapin *et al*., 1992), and adipocyte lipid-binding protein (ALBP) (Xu *et al*., 1992, 1993). Second, increased access to the ligand-binding pocket is either not observed in the apo-structure (I-FABP (Scapin *et al*., 1992), CRBPII (Winter *et al*., 1993), and ALBP (Xu *et al*., 1993)), or due to the formation of an intermolecular β-sheet in the crystal lattice (CRABPI; Thompson *et al*., 1995). Therefore the comparisons between the corresponding apo- and holo-structures give limited information on the flexibility of the portal components or on ligand-induced conformational changes.

The crystal structure of holo-CRABPII, as a CRABPII·all-*trans*-RA complex, was determined at 1.9 Å, where the methylene groups of the side-chain of Arg111 are in the center of CRABPII's hydrophobic core, and the guanidino group interacts with the carboxylate group of RA *via* a water molecule (Kleywegt *et al*., 1994). The R111M mutant of CRABPII is more than 40 times lower in affinity for all-*trans*-RA compared with the wild-type protein; nevertheless, it is properly folded, and its conformation is very similar to that of wild-type CRABPII (Wang *et al*., 1997). Here, we report the crystal structure of human apo-CRABPII (R111M) at 2.3 Å resolution, where we observed for the first time a portal conformation that allows RA to enter the binding pocket. The interesting features revealed from this apo-structure, and the results of crystal structure-based molecular modeling studies have led to the postulation of a three-step mechanism of RA entry.

Results and Discussion

Overall structure and crystal packing

The present mutant structure at 2.3 Å resolution is the first unliganded crystal structure of CRABPII. The structure, containing two independent molecules, Mol A and Mol B (Figure 1), and 287 water molecules, was refined to a crystallographic *R*-factor of 0.18 with good geometry (Table 1). As assessed by the program Procheck (Laskowski, 1993), the stereochemical parameters of the struc-

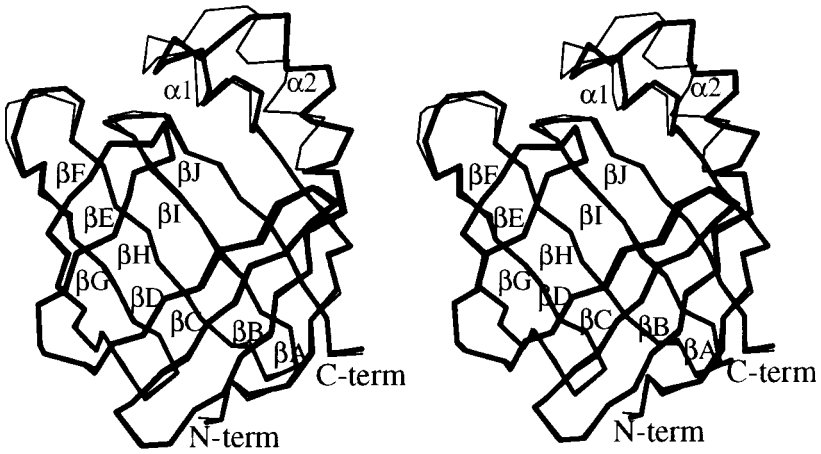

Figure 1. Stereodiagrams of apo-CRABPII (R111M). The C^{α} traces of Mol A (thick lines) and Mol B (thin lines) resulting from the structure-based alignment (see the text). The β-strands A through J and α-helices 1 and 2 are labeled. The illustration was prepared with the program Bobscript (R. Esnouf, robert@s151h16.rega.kuleuven.ac.be), an extended version of Molscript (Kraulis, 1991).

Table 1. Crystallographic data and refinement results of apo-CRABPII (R111M)

Resolution (Å)	2.3
No. reflections collected	32,588
No. unique reflections	10,509
R_{sym}	0.059
Data completeness (%)	85.4
Last shell (2.3 to 2.4 Å) completeness (%)	58.3
Space group	*P*1
Unit cell parameters	
a, *b*, *c* (Å)	37.58, 62.27, 35.12
α, β, γ (°)	106.40, 90.69, 110.63
No. independent molecules	2
Data cutoff in refinement	2σ(*F*)
R-factor	0.18
Number of water molecules	287
Total number of atoms	2463
rms deviation from ideal geometry	
Bond length (Å)	0.01
Bond angle (°)	1.24
Dihedral (°)	27.8
Improper angle (°)	1.06
Average temperature factor (Å^2)	
Protein atoms	37.7
Water oxygen atoms	50.1

ture are within one standard deviation of the mean for structures at comparable resolutions, or better. A few residues are in generously allowed regions (Val58A, Asp126B, Glu73A, and Glu73B) or disallowed regions (Asp126A and Met27A). Except for Glu73A and Glu73B, these residues are located in sharp turns between antiparallel β-strands or between antiparallel helices. In all cases, the electron density is well defined. The occurrence of such conformations in similar positions has been observed in other retinoid-binding proteins (Kleywegt *et al*., 1994; Thompson *et al*., 1995). Residues Glu73A and Glu73B are located at the center of βE. The deviation of their conformations from commonly observed ones might be a result of the hydrogen-bonding interactions of their side-chains with tryptophan side-chains (Trp109A and Trp109B) in the neighboring βH.

Mol A and Mol B, both bearing global resemblance to holo-CRABPII (Kleywegt *et al*., 1994), are found in two distinctly different crystal-packing environments. In particular, the α2 helix of Mol A is not involved in any intermolecular hydrogen bond (Table 2) and is exposed to solvent channels in the crystal lattice, whereas that of Mol B is tightly packed against the hairpin βI-βJ of Mol A. Consequently, the α2 helix of Mol B assumes a conformation significantly different from that of Mol A (Figure 1).

Conformation of Mol A is compatible with ligand entry

As shown in Figure 1, the most striking difference between Mol A and Mol B occurs at the α2 helix. We noticed at the early stage of the refinement that the α2 helix of Mol A had very poor electron density. We therefore refined the structure without including this helix until it was unambiguously defined by difference Fourier maps (see Materials and Methods). The well-defined final $2F_o - F_c$ electron density map for residues Leu18 to Ala40 in Mol A and Mol B are shown in Figure 2. The omit map (Bhat, 1988) has the same appearance as the final $2F_o - F_c$ map. As expected, the α2 helix of Mol A has a higher than average temperature factor in the final structure. This observation is compatible with the notion that this structural element is intrinsically flexible and closely related to the mechanism of ligand entry.

The differences between the α2 helices of Mol A and Mol B stand out clearly from other structural motifs in the entire molecule. The α2 helix of Mol B is straight and has three complete helical turns, whereas that of Mol A partially loses its integrity (Figure 1). The N terminus of the α2 helix in Mol A is unwound. According to the secondary-structure analysis (Kabsch & Sander, 1983), the α2 helix of Mol B starts at residue 25, whereas that of Mol A starts at residue 27. Both helices end at residue 36.

Table 2. Intermolecular interactions of the α2 helix in eight apo-iLBP crystal structures

iLBP	PDB code	Space group	Resolution (Å)		H-bond[a]	vdw contact[b]	Reference
CRABPII (R111M)	1xca	*P*1	2.3	Mol A	0	8	This work
				Mol B	7	43	
CRABPI	1cbi	$P3_121$	2.7	Mol A	1	22	Thompson *et al*. (1995)
				Mol B	3	20	
CRBPII	1opa	*P*1	1.9	Mol α	5	33	Winter *et al*. (1993)
				Mol β	8	43	
I-FABP	1ifb	$P2_1$ form1	1.96		3	26	Sacchettini *et al*. (1989b)
I-FABP	1ifc	$P2_1$ form2	1.19		8	62	Sacchettini *et al*. (1989a)
M-FABP[c]	1ftp	$P2_1$	2.2	Mol A	2	28	Haunerland *et al*. (1994)
				Mol B	2	26	
ALBP	1alb	$P2_12_12_1$	2.5		3	29	Xu *et al*. (1992)
ALBP	1lib	*C*2	1.7		2	21	Xu *et al*. (1993)

Residue range of 25 to 37 was used as the α2 helix region, except for I-FABPs, where a range of 24 to 34 was used.
[a] ⩽ 3.5 Å.
[b] ⩽ 4.0 Å.
[c] Muscle fatty acid-binding protein.

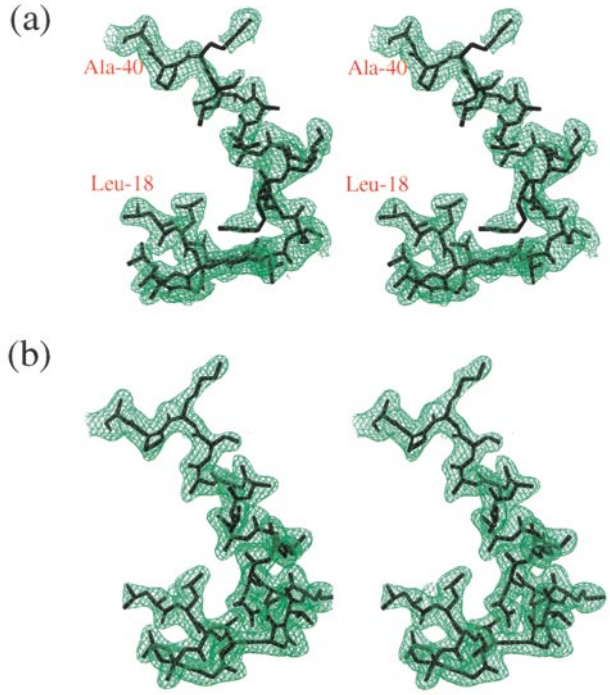

Figure 2. The $2F_o - F_c$ electron density map contoured at 1σ level in the regions around the α2 helices in Mol A (a) and Mol B (b) of apo-CRABPII (R111 M) with the final model superimposed. The representation was prepared with the program Bobscript (see the legend to Figure 1 for details).

The deviations between the corresponding C^α positions become significant from the C terminus of α1 and residues between the α1 and α2 helices. This trend continues into the α2 helix and culminates at residues 28 to 31 with deviations of 3.5 to 4.5 Å. The deviations then subside in the loop between α2 and βB. The overall rms deviation between all C^α atoms of Mol A and Mol B is 1.22 Å. However, if residues 22 to 38 (the α1-loop-α2 motif) are excluded from the calculation, the rms deviation drops sharply to 0.59 Å. For residues 22 to 38, the rms deviation is 3.1 Å. The conformational difference in the α2 helix extends over the loops and secondary structure elements before and after the helix. For example, the deviation between the corresponding C^α positions of Gly23 (one of the loop residues between the two helices) is 2.5 Å. In addition, the βG-βH hairpin loop adopts different conformations in the two molecules (Figure 1).

The conformational differences in the α2 region and in the βG-βH region between Mol A and Mol B are the results of the dissimilarities in their crystal-packing environment. The α2 helix of Mol B is involved in seven intermolecular hydrogen bonds and 43 intermolecular van der Waals (vdw) contacts in the crystal lattice, whereas that of Mol A is free from intermolecular hydrogen bonding, with only eight intermolecular vdw contacts (Table 2). Among the eight vdw contacts, five involve Val26,

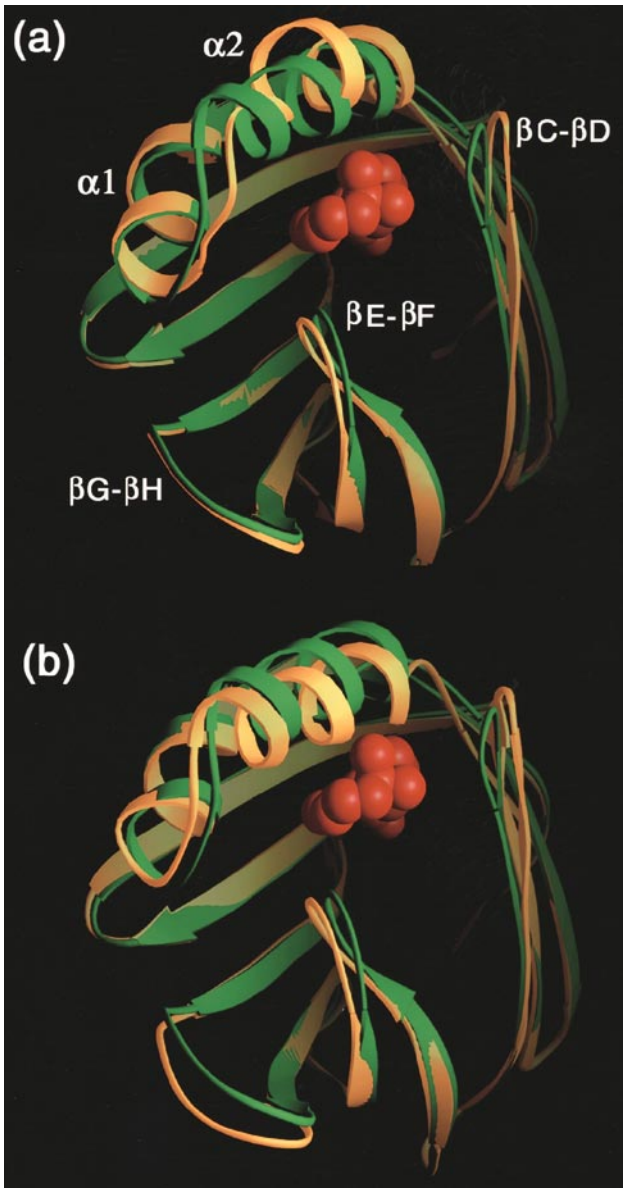

Figure 3. Ribbon representations of the two apo-CRABPII (R111 M) molecules (in yellow), Mol A (a) and Mol B (b). The CRABPII·RA complex (protein in green and RA in red) (Kleywegt *et al.*, 1994) is superimposed with Mol A and Mol B according to the structure-based alignment. The view is along the long axis of RA looking into the binding cavity. Several structural elements are labeled in (a). The illustration was prepared with the program Ribbons (Carson, 1987).

which is part of the loop between the α1 and α2 helices, and three are contacts with atom NZ of symmetry-related Lys101, whose side-chain is almost always flexible. Therefore, these eight vdw contacts do not have a significant impact on the conformation of the α2 helix in Mol A. Figure 3 illustrates the structure-based superposition of Mol A and Mol B, respectively, with the CRABPII·all-*trans*-RA complex (Kleywegt *et al.*, 1994). The α2 helix in Mol A unwinds at the N terminus and moves outward from the RA-binding pocket

(Figure 3(a)). On the contrary, the α2 helix of Mol B is compressed toward the binding pocket (Figure 3(b)). There are three consequences of this conformational change in Mol B. First, the RA entrance of the binding pocket closes up. Second, there is not enough room for RA in the pocket. Third, the βG-βH loop moves concurrently with the α2 helix by approximately 2.6 Å (Figure 3(b)), which is not observed either in Mol A or in holo-CRABPII (Figure 3(a)). Therefore, the conformation of Mol B appears to be an artifact of crystal packing, which appears to be responsible for the fact that the unique conformation of α2, as seen in Mol A, was not previously observed in the crystal structures of other apo-iLBPs.

There are seven previously determined crystal structures of apo-iLBPs in the PDB. Examination of the seven structures reveals that, without exception, the α2 helix is involved in intermolecular hydrogen bond(s) and extensive intermolecular vdw contacts (Table 2). Among the seven apo-iLBP structures, the α2 helix in the 1.19 Å apo-I-FABP structure (Scapin *et al.*, 1992) is involved in eight intermolecular hydrogen bonds and 62 intermolecular vdw contacts (Table 2). These interactions, especially the electrostatic ones, must have significant stabilizing effect on the α2 helix. Among all apo-iLBP structures, MolA of apo-CRABPII (R111M) shows its uniqueness in crystal packing where the α2 helix is not involved in any intermolecular electrostatic interactions. In the following sections, apo-CRABPII (R111M) refers to Mol A, if not otherwise specified.

Comparison of apo-CRABPII (R111M) with holo-CRABPII

Unique portal conformation

The differences between the apo-CRABPII (R111M) (this work) and holo-CRABPII (Kleywegt *et al.*, 1994) occur mostly at the end of the β-barrel where RA appears to enter. When viewed along the long axis of bound RA, differences are clearly seen in the region of the α2 helix and the two nearby hairpin loops, βC-βD and βE-βF (Figure 3(a)). Quantitative illustrations of the distribution of C^{α} deviations are shown in Figure 4(a), where the differences in the α2 helix region are the largest. Compared with that of holo-CRABPII, the α2 helix of apo-CRABPII (R111M) is partially unwound and displaced (Figure 3(a)). The conformational changes of the βC-βD and βE-βF hairpin loops in the apo-molecule are concerted with that of the α2 helix. Both loops move outward, so that the opening of the ligand-binding pocket is wide enough for RA to enter (Figure 3(a)). The largest deviation of the C^{α} atoms is at Val58, the tip of the βC-βD loop, and reaches 3.3 Å (Figure 4(a)). The large scope of the movement of this hairpin is made possible by the lack of hydrogen bonding between β-strands D and E.

Both the apo- and holo-structures have been determined for two more retinoid-binding proteins, CRABPI (Kleywegt *et al.*, 1994; Thompson *et al.*, 1995) and CRBPII (Winter *et al.*, 1993). All four structures have more than one independent molecules in the asymmetric unit. Therefore, an apo-structure may be compared with different holo-structures, and *vice versa*. Although the rms deviation varies, some pairs exhibit almost identical conformations within experimental error (Winter *et al.*, 1993). Illustrated in Figure 4(b) and (c) are deviations of C^{α} atoms between corresponding apo- and holo-forms of CRABPI and CRBPII. CRBPII has two copies of the molecule in the apo-structure and four copies in the holo-structure, whereas CRABPI has two copies in both apo- and holo-structures. Results are plotted for the pairs of molecules with the largest rms deviations (Figure 4(b) and (c)). apo- and holo-CRABPIs display the largest deviations in the βC-βD hairpin region, which is the consequence of the formation of an intermolecular β-sheet in the crystalline state (Thompson *et al.*, 1995). Although some conformational difference is evident in the loop between α2 and βB, no significant conformational change can be seen in the α2 helix region (Figure 4(b)). apo- and holo-CRBPIIs show obvious conformational variations in the C-terminal half of the α2 helix, but the βC-βD hairpin region does not show any significant deviation (Figure 4(c)). Only between apo-CRABPII (R111M) and holo-CRABPII are substantial differences found in all three regions: the α2 helix, and the βC-βD and βE-βF loops. The difference in the α2 helix extends over the entire region of loop-α2-loop even into the α1 helix and the βB strand (Figure 4(a)).

The binding of RA to CRABPs and that of lipid-like ligands to iLBPs are characterized by the deep localization of the ligand inside the β-barrel cavity. Although the binding specificity is influenced by several deeply buried amino acid residues that interact with the carboxylate of the ligand, the majority of the ligand–protein contacts are provided by side-chains that are not localized deep in the barrel. Figure 4(d) illustrates the solvent accessibility changes of amino acid residues upon removal of the ligand in the CRABPII·RA complex (Kleywegt *et al.*, 1994). The values may also represent the contact surface that each amino acid provides to bound RA. Since hydrophobic interactions that contribute significantly to the binding of hydrophobic ligands to their cellular host proteins, the contact area should be closely related to the binding affinity. Three structural motifs, the α2 helix and the βC-βD and βE-βF loops, account for ~80% of the total contact area. It is, therefore, not a coincidence that the conformation of these structural motifs undergoes the most significant changes before and after ligand binding, as observed in apo-CRABPII (R111M) (Figures 3(a) and 4(a)).

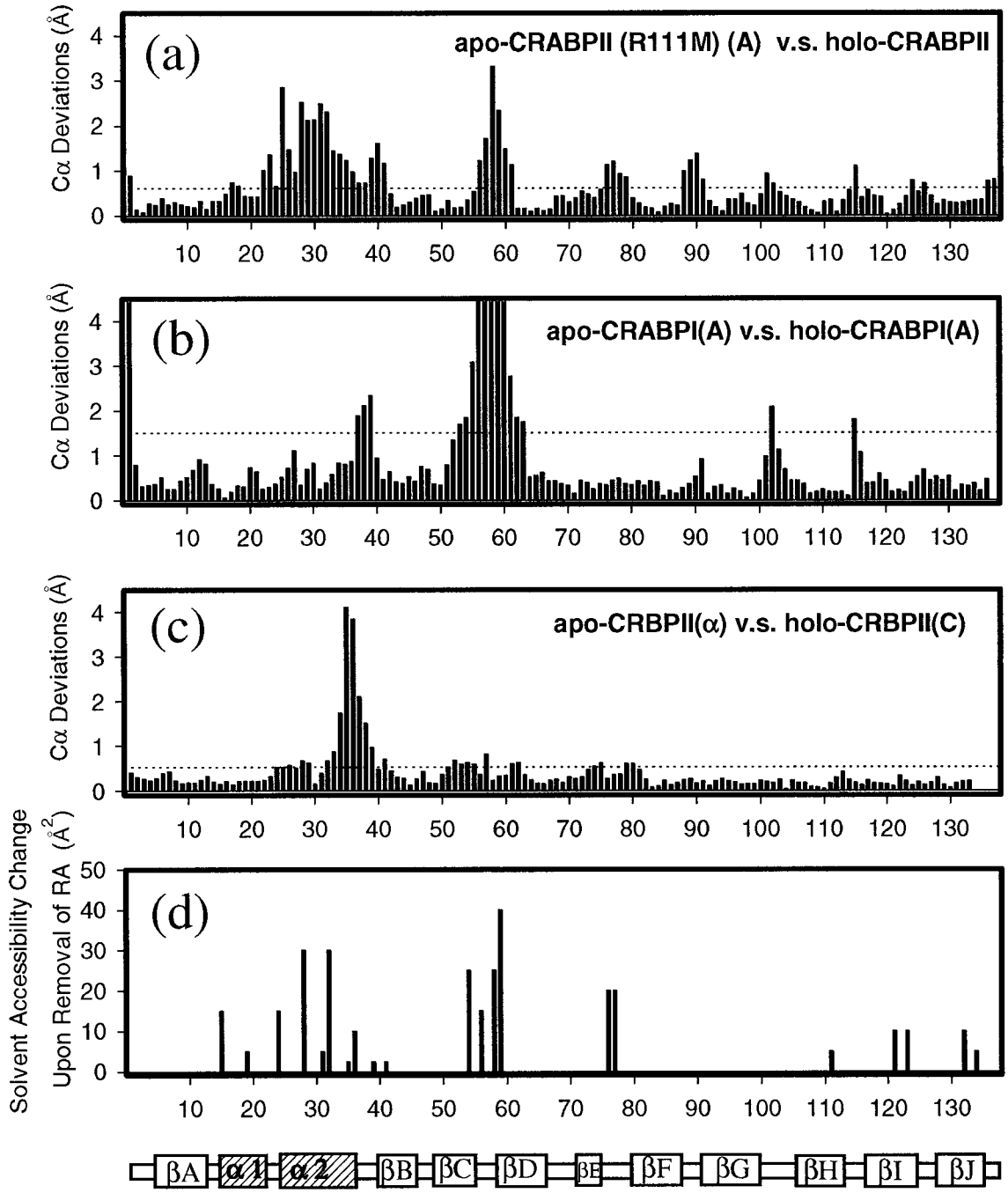

Figure 4. C^{α} atom deviations (a) between Mol A of apo-CRABPII (R111 M) (this work) and holo-CRABPII (Kleywegt *et al.*, 1994), (b) between apo- (Thompson *et al.*, 1995) and holo-CRABPI (Kleywegt *et al.*, 1994) (copy A of both apo- and holo-structures is used), and (c) between apo- and holo-CRBPII (Winter *et al.*, 1993) (copy α of apo- and copy C of holo-CRBPII are used). The broken lines indicate the standard deviation. The change in solvent accessibility of amino acid residues upon removal of RA in the CRABPII·RA complex (Kleywegt *et al.*, 1994) is illustrated in (d). A secondary-structure assignment is shown at the bottom.

Side-chains and solvent

Twenty-one amino acid residues have been shown to be in contact with the bound RA in holo-CRABPII, among which 18 residues change their accessible surface area by >1 Å^2 upon removal of RA (Figure 4(d)) (Kleywegt *et al.*, 1994). A comparison of these 21 residues in the two structures reveals that most of these side-chains in apo-CRABPII (R111M) retain similar positions and/or conformations as in holo-CRABPII, except for residues Ile28, Ile31, Ala32, Val58, Arg59, and Arg132. The first three residues are located in the α2 helix, and Val58 and Arg59 are in the βC-βD hairpin. Both α2 and βC-βD have significantly different conformations with and without bound RA. The changes in these side-chains are as expected. The

side-chain of Arg132 forms a hydrogen bond with the carboxylate of RA in holo-CRABPII (Kleywegt *et al.*, 1994) and, therefore, is expected to undergo a conformational change when RA is not present. In addition to these RA-contacting side-chains, the side-chain of Arg29 on the α2 helix shows significant conformational differences between apo-CRABPII (R111M) and holo-CRABPII. Illustration of the binding pocket are shown in Figure 5. Unlike smaller side-chains of Ile28, Ile31, Ala32 and Val58, those of Arg29, Arg59, and Arg132 show dramatic differences between the apo- and holo-structures.

In apo-CRABPII (R111M), the guanidino group of Arg29 points into the binding pocket and forms a hydrogen bond with the carbonyl group of Asn25; and the side-chain of Arg59 extends directly into the cavity and participates in the formation of a hydrogen bond network with Arg132, Tyr134, and five water molecules. The shortest distance between the guanidino group of Arg59 and that of Arg132 is 6.20 Å. Two water molecules bridge the interaction between the two guanidino groups (Figure 5(a)). With the bridging water molecules, the repulsive positive charges of the two arginine side-chains may be stabilized. The guanidino group of Arg29 is approximately 11 Å away from those of Arg59 and Arg132. The three arginine side-chains provide positive charges extending from the entrance to the middle of the RA-binding pocket (Figure 5(a)).

In holo-CRABPII (Kleywegt *et al.*, 1994), all three arginine side-chains have significantly different conformations (Figure 5(b) and (c)). Arg29 on the α2 helix swings its guanidino group out of RA-binding pocket and forms a salt bridge with Glu16 that is on the α1 helix. Arg59 on βC-βD forms a hydrogen bond with Gln74, which is located on the βE-βF loop. With this conformation, Arg59 is able to provide hydrophobic surface contact with bound RA by extending the aliphatic chain across the hydrophobic isoprene and ionone ring system. Arg132 forms hydrogen bonds with the carboxylate of RA and also with the carbonyl group of Ala36 that is located at the C terminus of the α2 helix. It is noticeable that the electrostatic interactions involving the three arginine side-chains may help to stabilize the complex by locking up all three flexible structural elements at the portal, the α2 helix and the βC-βD and βE-βF hairpin loops (Figure 5(b)).

The side-chain conformational changes of Arg29, Arg59, and Arg132 are associated with the conformational flexibility of the α2-helix and the βC-βD and βE-βF hairpin loops. In particular, the conformations of Arg29 and Arg132 observed in apo-CRABPII (R111M) are related to the unwinding and displacement of the α2 helix. As expected, they are not seen in either Mol B (apo-CRABPII (R111M)) or holo-CRABPII. The side-chain conformation of Arg59 is closely related to the conformational changes of the βC-βD and βE-βF loops. When both hairpin loops move outwards to a certain extent, the hydrogen bond between the sides chains of Arg59 (on βC-βD) and Gln74 (on βE-βF) will be broken. The Arg59 side-chain is able to assume the conformation as seen in apo-CRABPII (R111M) (Figure 5(a)). This conformation is observed in both Mol A and Mol B (apo-CRABPII (R111M)) because the two hairpin loops assume identical conformations in the two molecules (Figure 1). In contrast, the structure-based alignment of the two independent molecules of apo-CRABPI with holo-CRABPI reveals that the βE-βF hairpin loop has identical conformations in all three molecules, whereas the βC-βD loop assumes three different conformations (Figure 2 of Thompson *et al.*, 1995). In the two apo-CRABPI molecules, the βC-βD loop moves outwards to different extents probably due to the formation of the intermolecular β-sheet between the two βD strands. In molecule B, the conformational change is minor, and the salt bridge (3.26 Å) between Arg59 and Glu74 (Gln74 in CRABPII) prohibits the side-chain of Arg59 from pointing into the RA-binding pocket. In molecule A, although the conformational change in the βC-βD loop breaks the salt bridge between Arg59 and Glu74, the guanidino group of Arg59 is involved in four intermolecular electrostatic interactions (2.85 to 3.38 Å) with molecule B residues, and it is consequently impossible for it to point into the RA-binding pocket, either.

Twelve water molecules are found in the RA-binding pocket of apo-CRABPII (R111M) (this work), as shown in Figure 5(a). Fourteen are found in holo-CRABPII (Kleywegt *et al.*, 1994), as shown in Figure 5(b). The R111M mutation appears to have eliminated two water molecules associated with the guanidino group of Arg111 (Figure 5(c)). The α2 helix faces the RA-binding pocket with short hydrophobic aliphatic amino acid side-chains, and the entrance of the pocket is surrounded by hydrophobic residues. The pocket is, therefore, not likely to house more water molecules in apo-CRABPII (R111M). With a larger volume of the binding pocket the water molecules in apo-CRABPII (R111M) appear to be less organized and to be loosely packed. Figure 5(c) indicates that the binding of RA displaces six water molecules. However, they are not expelled from the binding pocket. Instead, they are squeezed into the inner portion of the pocket, filling the space between RA and the protein with the formation of a much more compact hydrogen bond network (Figure 5(b)). The water structure in the RA-binding pocket appears to be a ''cushion'' that fills the pocket in the absence of ligand and occupies a smaller space when RA is bound. Upon ligand entry, RA seems to be guided into the pocket, with its smaller and negatively charged carboxylate end first. Consequently, the larger isoprene and ionone ring system of RA may block the entrance, leaving no path for the pocket water molecules to escape.

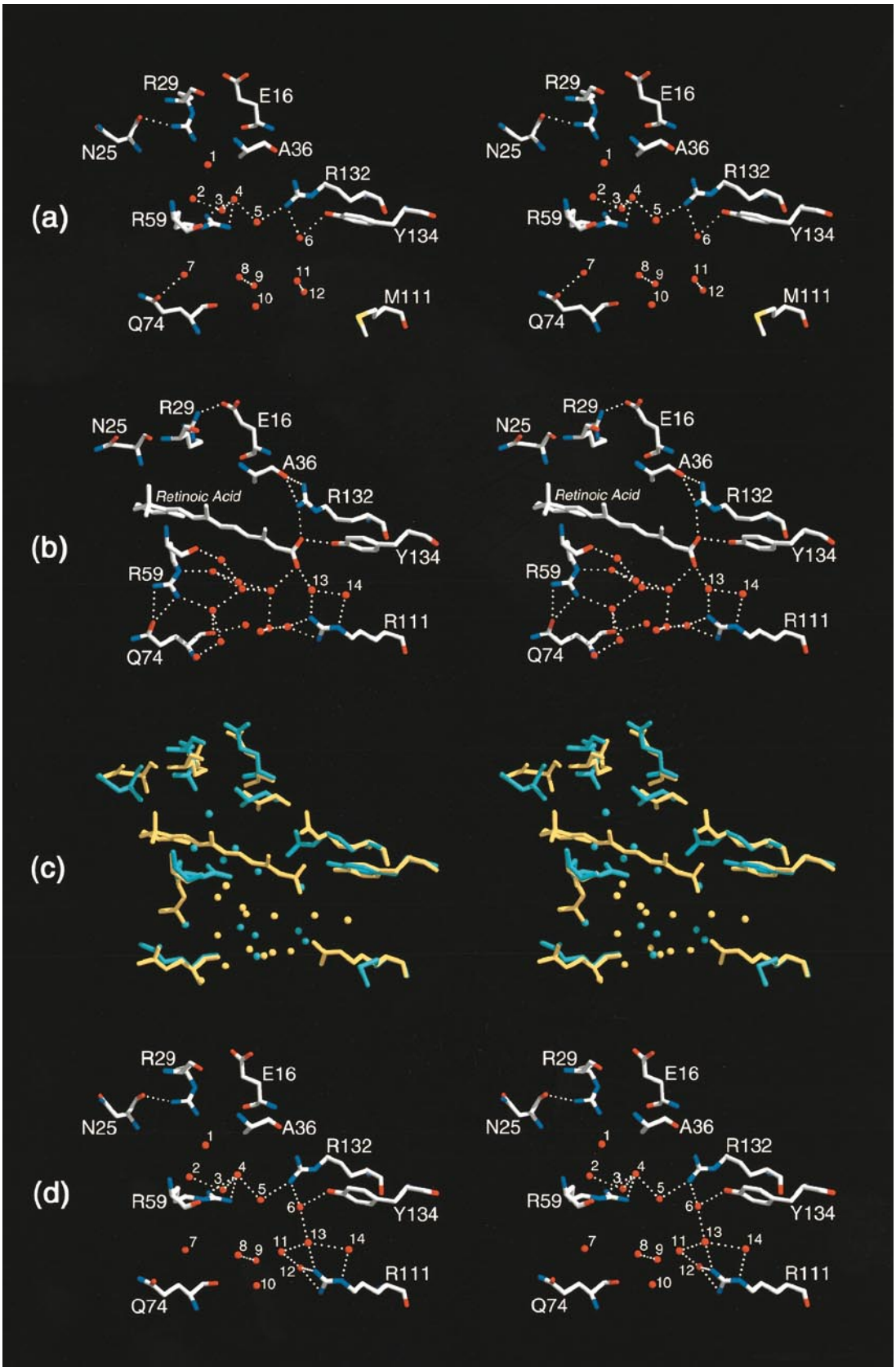

Figure 5. Stereodiagrams of the RA-binding pocket region of (a) apo-CRABPII (R111M) (Mol A, this work), (b) holo-CRABPII (Kleywegt *et al*., 1994), (c) superposition of apo-CRABPII (R111M) (in blue) and the holo-protein with all-*trans*-RA molecule (in yellow), and (d) a model of wild-type apo-CRABPII. The protein residues are illustrated as rods, and water molecules as balls, with carbon in white, nitrogen in blue, oxygen in red, and sulfur in yellow. Labels are provided in (a), (b), and (d). Hydrogen bonds are shown as dotted lines and are omitted in (c) for clarity. The illustration was prepared with the program TkRaster3D (Hillary Gilson, National Institute of Standards and Technology, http://indigo15.carb.nist.gov/TkRaster3D), a version of Raster3D (Bacon & Anderson, 1988) with enhanced features.

Possible impact of R111M mutation on apo-CRABPII structure

Arg111 is in the hydrophobic core of CRABPII. The impact of R111M mutation on apo-CRABPII structure can be assessed in three aspects. First, the mutation should not weaken the hydrophobic core of the molecule and therefore should not cause a dramatic change in the overall fold. This has been demonstrated from solution NMR studies (Wang *et al*., 1997). Second, such alteration is not expected to have much impact on the distal conformation of the portal that is more than 15 Å away. This is demonstrated by the distinct conformation of the two α2 helices in the two crystallographically independent molecules bearing the same mutation in the present structure. Third, the elimination of the guanidino group, however, removes one positive charge from the binding pocket, and alters the solvent structure and the hydrogen bond network in the pocket (Figure 5(a) and (b)). To evaluate the consequences of the structural perturbation, an apo-CRABPII molecule was built by using a molecular modeling strategy applicable to slightly modified crystal structures (see Materials and Methods). The energy-minimized model of apo-CRABPII is shown in Figure 5(d). Four guanidino groups coexist in the RA-binding pocket with that of Arg29 at the portal, those of Arg59 and Arg132 in the middle, and that of Arg111 at the bottom. Since NH2 is the closest atom in between the four arginine side-chains, the distances mentioned below are those between the corresponding NH2 atoms. In the crystal structure of apo-CRABPII (R111M), the distances from Arg132 to Arg59 and Arg29 are 6.20 and 11.44 Å, respectively (Figure 5(a)). In the model of apo-CRABPII (Figure 5(d)), the distances from Arg132 to Arg59, Arg29 and Arg111 are 6.37, 11.68, and 7.96 Å, and the distances from Arg-111 to Arg-59 and Arg-29 are 9.10 and 18.65 Å, respectively. It is therefore suggested that the repulsion between two guanidino groups 6 to 8 Å apart could be attenuated by two bridging water molecules, and that the repulsion might not be significant when the distance in between reaches 9 Å. The most significant difference between apo-CRABPII (R111M) and apo-CRABPII (model) is the hydrogen bond network in the RA-binding pocket (Figure 5(a) and (d)). Although the difference in terms of the number of water molecules is only two, the difference in terms of the number of hydrogen bonds is about 50%. In particular, in apo-CRABPII (R111M), the side-chain of Arg59 is involved in a hydrogen bond network with Arg132, Tyr134, and five water molecules (Figure 5(a)). The network is extended, in apo-CRABPII (model), to Arg111 and six more water molecules (Figure 5(d)). Since R111 interacts with bound RA *via* the hydrogen bond network (Figure 5(b)), we suggest that the reduction in the hydrogen bond network due to R111M mutation is responsible for the low RA affinity of the mutant protein.

A three-step mechanism of ligand entry

Ligand entry through the crevice formed by the α2 helix and the βC-βD and βE-βF loops has been suspected from previous studies (Banaszak *et al*., 1994; Winter *et al*., 1993), because this region is the only site where the end of the bound ligand (RA, retinol, lipid, or fatty acid) extends to the surface of the protein. Before this study, however, there was no unliganded crystal structure compatible with this view of ligand entry (for references, see Table 2). Recently, backbone disorder and increased mobility of the α2 helix and the βC-βD loop were suggested by solution NMR studies of CRABPII (Rizo *et al*., 1994), where it was shown that a region around the C terminus of the α2 helix was much more flexible in the apo-protein. Although the flexibility of the βC-βD loop was also shown, that of the βE-βF loop was not suggested (Rizo *et al*., 1994). More recently, the solution structure of apo-I-FABP was determined by three-dimensional NMR methods (Hodsdon & Cistola, 1997a), and the backbone dynamics of both apo- and holo-I-FABPs were characterized and compared by using amide ^{15}N relaxation and ^{1}H exchange NMR measurements (Hodsdon & Cistola, 1997b). In addition to the residues in the vicinity of the α2 helix and the βC-βD loop, the residues of the βE-βF loop were also found more mobile than other parts of the molecule. Although both studies suggest that increased mobility and discrete disorder in the α2 helix (C-terminal half) and the βC-βD hairpin loop permit ligand entry into the binding cavity, the conformational change is hard to envision and the proposed mechanism of portal opening appears contradictory. Rizo and coworkers (1994) suggested that the region around the C terminus of helix α2 may act as a flexible hinge to facilitate opening of the portal, in support of the view of a hinged rotation of the entire α1-loop-α2 motif relative to the rest of the protein as suggested by limited proteolysis experiments (Jamison *et al*., 1994). On the contrary, Hodsdon & Cistola (1997a) proposed an order–disorder transition with a fraying of the distal half of the α2 helix and a decoupling of long-range interactions with the βC-βD loop. The two views were completely different and the conformational change in the βE-βF loop was either not observed (Rizo *et al*., 1994) or not obviously related to the proposed mechanism (Hodsdon & Cistola, 1997b). In the crystal structure of apo-CRABPII (R111M), we have observed a concerted conformational change in the α2-helix and the βC-βD and βE-βF hairpin loops, which is compatible with RA entry (Figure 3(a)).

A few years ago, based on the calculation of electrostatic potential of a CRABP model structure, Zhang *et al*. (1992) proposed that the entry of RA into the binding pocket might be guided by a positive potential, extending from the cavity to the entry site. Although the homology modeling at the time failed to describe accurately the details of

the structure (Kleywegt *et al.*, 1994), this RA-guiding potential appears to have found experimental support from the current structure of apo-CRABPII (R111M), where Arg29, Arg59 and Arg132 provide positive charges from the portal to the middle of the pocket. Considering also Arg111, as suggested by our molecular modeling studies, the positive potential seems indeed to guide the RA molecule all the way into the binding pocket (Figures 5(d) and 6(a)), although it is not conclusive, until an apo-CRABPII structure compatible with RA entry is available.

The partial unwinding of the α2 helix, as observed in our crystal structure, assists the ligand entry in two ways. First, the in-and-out movement of the α2 helix would efficiently open the binding pocket, because the bound ligand lies ''underneath'' and along the α2 helix (Kleywegt *et al.*, 1994). Second, the unwinding at the N terminus of the α2 helix appears to be crucial for the helix to move outwards to the extent that allows RA to enter. In contrast to the tight fit of RA bound in the CRABPII·RA complex (Figure 6(b)), the position of RA in the cavity of Mol A could almost allow its free movement in and out (Figure 6(a)).

In summary, we propose a three-step mechanism of RA entry, i.e. opening, directing, and locking. First, the opening of the RA-binding pocket is made possible by a concerted conformational change in three structural elements: the α2 helix and the βC-βD and βE-βF hairpin loops, where the unwinding of the α2 helix at its N terminus is essential (Figure 3(a)). Second, closely associated with the conformational changes in the three structural elements, three arginine (Arg29, Arg59, and Arg132) side-chains, point into the RA-binding pocket; and the resulting electrostatic potential appears to be responsible for directing entry of RA into the binding pocket (Figures 5(d) and 6(a)). Third, upon RA entry, the guanidino groups of Arg29, Arg59, and Arg132 move away from RA and form hydrogen bonds and/or salt bridges, which seem to lock all three flexible structural elements, and therefore to stabilize the protein·RA complex (Figures 5(b) and 6(b)).

Materials and Methods

Protein expression and purification

Protein expression and purification procedures reported by Wang *et al.* (1997) were followed and are described briefly here. Recombinant human CRABPII was expressed from bacterial strain BL21(DE3)pLysS transformed with expression vector pET-17b containing CRABPII cDNA. Site-directed mutagenesis of the R111M mutant was performed according to the method of Kunkel (1985). Four liters of LB medium containing ampicillin (100 μg/ml) and chloramphenicol (25 μg/ml) was inoculated with BL21(DE3)pLysS/pET-17b/CRABPII (R111M), followed by incubation at 37°C until the A_{600} reached 0.6 to 0.8. Isopropyl-1-thio-β-D-galactopyranoside was then added to a final concentration of 0.4 mM. After a further three to five hour of growth at 30°C, the bacterial cells were harvested by centrifugation and suspended in 100 ml of 10 mM Tris-HCl (pH 8.2) with 1 mM dithiothreitol. After sonication and centrifugation, the supernatant was applied to a DEAE-cellulose column and washed with the same buffer. Eluate fractions from the linear NaCl gradient (0 to 200 mM) were monitored at A_{280} and collected, followed by analysis by 15% SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing CRABPII were concentrated using an Amicon ultrafiltration cell with a YM 10 membrane, and further separated with a Sephadex G-50 column. The fraction containing CRABPII was assessed by 15% SDS- PAGE to be more than 99% pure. The average yield of the expression was about 35 mg/l.

Crystallization and data collection

Although many attempts to crystallize apo-CRABPII had been unsuccessful, crystals of apo-CRABPII (R111M) were obtained by using the hanging-drop crystallization method. Crystals grew in droplets of 15 mg/ml protein in 100 mM Tris-HCl (pH 8.0) with 100 mM sodium acetate in 15% polyethylene glycol (PEG) 8000 precipitant, equilibrated against 100 mM Tris-HCl (pH 8.0) with 200 mM sodium acetate and 30% PEG 8000. A crystal of 1.0 mm × 0.3 mm × 0.2 mm was used for the entire data collection at room temperature. A Siemens multiwire area detector mounted on a Rigaku rotating anode X-ray generator operated at 40 kV and 100 mA was used with graphite-monochromatized Cu Kα radiation (λ = 1.5418 Å). The crystal-to-detector distance was

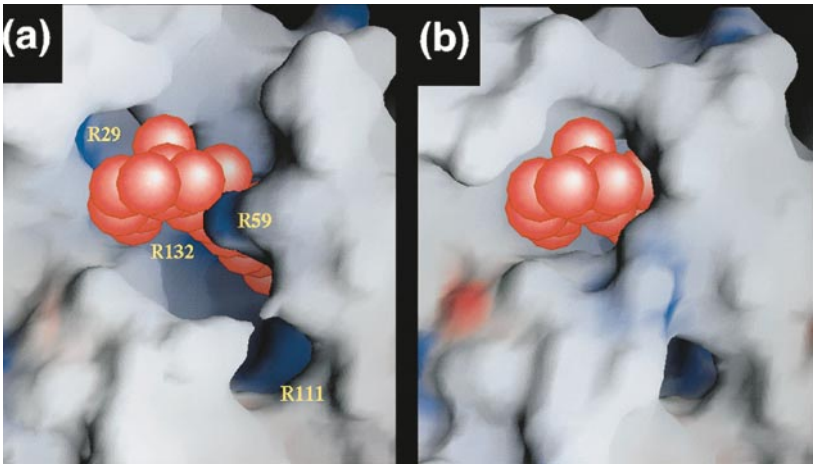

Figure 6. The vdw surface representations of CRABPII. (a) The accessibility of the binding pocket in apo-CRABPII (model, see the text) to the ligand. RA (in red) is shown as a reference, based on the superposition of apo-CRABPII with the CRABPII·RA complex (Kleywegt *et al.*, 1994). The tight binding of RA in holo-CRABPII is illustrated in (b) for comparison. By electrostatic potential, the vdw surface is color coded as blue for positive and red for negative. Arginine residues 29, 59, 111, and 132 are labeled in (a). The illustration was prepared with the program Grasp (Nicholls *et al.*, 1991).

10 cm. The data set was collected by using an ω-scan in 0.25° steps with an X-ray exposure time of 240 seconds per frame. The crystal did not show any sign of significant radiation damage. The data set was processed using XENGEN (Howard *et al.*, 1987) v1.3 to 2.3 Å with an R_{sym} of 0.059. The overall $I/\sigma(I)$ was 10.4 and that for the last shell (2.30 to 2.31 Å) was 1.29. Crystallographic data are summarized in Table 1. The crystal is triclinic, in space group *P*1, with unique cell dimensions not seen before for iLBPs. The unit cell volume includes two independent molecules.

Structural solution and refinement

The structure was solved by molecular replacement techniques with the program AmoRe (Navaza, 1994) using the structure of the CRABPII·RA complex (Kleywegt *et al.*, 1994) with RA and water molecules removed as the search model. X-ray diffraction data in the range of 10 to 4 Å were used. The correlation coefficients for the top five peaks from the rotation search were 0.385, 0.324, 0.198, 0.152 and 0.146. The first solution was found from the subsequent translational search. The top two outstanding peaks from the rotational search were the correct solutions as confirmed by the second translational search with the position of the first solution fixed. Two molecules were found for which the correlation coefficient was 0.60 and the crystallographic *R*-factor was 0.39. Rigid-body refinement within AmoRe improved the two values to 0.62 and 0.37, respectively (for 10 to 4 Å data). Global energy crystallographic refinement implemented in the program X-PLOR (Brünger, 1992) was used for the initial refinement of the structure. Randomly chosen 10% reflections were assigned for R_{free} calculations (Brünger, 1993) to monitor the refinement progress. The starting model with a uniformly assigned temperature factor of 25.0 Å^2 gave both *R*-factor and R_{free} values of 0.44 for $I/\sigma(I) \geqslant 1$ reflections in the range of 8 to 2.3 Å. First, rigid-body refinement was employed, with the two independent molecules treated as separate entities. Simulated annealing and positional refinement were then carried out with incrementally higher resolution limits starting at 3.0 Å. Grouped *B*-factor refinement was used until the resolution was extended to 2.3 Å. Several rounds of simulated annealing, positional refinement, and temperature-factor refinement lowered the *R*-factor and R_{free} to 0.25 and 0.34, respectively. At this stage, an omit map (Bhat, 1988) was calculated for every 15 residues. The entire model was checked and adjusted when necessary using the graphics programs Chain 5.0 (Sack, 1988) and O 5.10 (Jones *et al.*, 1991). We found that amino acid residues 24 to 37 in one of the two independent molecules (Mol A) had almost no $2F_o - F_c$ electron density at the 1σ level. The 14 residues of this segment were therefore excluded in the subsequent refinement to minimize the model bias and phase error. Non-crystallographic restraints between the two molecules were imposed except for the segment of residues 24 to 37. Weights of 300 and 100 were used for the restraints of the backbone and side-chain atoms, respectively. Further refinement lowered the R_{free} and *R*-factor to 0.31 and 0.24, respectively. The weights of the restraints were gradually reduced to zero within several rounds of refinement. Solvent molecules (water) were identified from positive peaks in the $F_o - F_c$ ($\geqslant 3\sigma$) maps and were required to be within the vicinity of at least one protein non-carbon atom or the oxygen atom of another water molecule ($\leqslant 3.5$ Å). The last recorded R_{free} was 0.282 for 10,509 reflections within 8 to 2.3 Å. However, the electron density for residues 24 to 37 of Mol A was still of insufficient quality to fit the missing segment. At this stage, the restrained least-squares refinement procedure (Hendrickson, 1985; Konnert, 1976) was employed with the program GPRLSA (Furey *et al.*, 1982). All data with $I \geqslant 1\sigma(I)$ and within 8.0 to 2.3 Å were included and geometry was tightly restrained. Several rounds of such refinement significantly improved the electron density of the missing segment of Mol A. Residues 24 to 32 were unambiguously built into the difference map. Further least-squares refinement allowed the addition of residues 33 to 37 to the model and improved the geometry of residues 24 to 32. The final model, with an *R*-factor of 0.18 for 10,509 reflections with $I \geqslant 1\sigma(I)$ and within 8.0 to 2.3 Å, contains 2,176 protein atoms of the two independent molecules and 287 oxygen atoms of water molecules. The coordinates have been deposited with the PDB under the accession code 1xca.

Molecular modeling of apo-CRABPII

Molecular modeling studies were carried out with X-PLOR 3.1 (Brünger, 1992) and O 5.10 (Jones *et al.*, 1991) on an SGI Indigo2 workstation with Solid Impact graphics. The initial model of apo-CRABPII was built based on the crystal structure of Mol A (apo-CRABPII (R111M)) where Met111 was replaced by Arg111 together with the two associated water molecules (nos 13 and 14 in Figure 5(b) and (d)) from holo-CRABPII (Kleywegt *et al.*, 1994). The initial model was subject to energy minimization and geometry optimization using the conjugate gradient method described by Powell (1977). All associated solvent molecules from the crystal structure were included in energy minimization and geometry optimization. The Engh & Huber (1991) geometric parameters were used as the basis of the force field. The total energy of the initial model was 488 kcal/mol and the energy minimization converged with a total energy of −937 kcal/mol, of which the electrostatic contribution was −589 kcal/mol. For comparison, the total energy and its electrostatic term of minimized Mol A structure were −936 and −582 kcal/mol, although its initial total energy was −8 kcal/mol. The rms deviation for all C^α atoms between the initial and optimized models was 0.24 Å. The shifts of ordered water molecules in the RA-binding pocket were as expected, so that the bad contacts between water molecules no. 11 and 12 with the guanidino group of Arg111 were removed (Figure 5(a) and (d)). The crystal structure-based energy minimization technique employed in this study was, therefore, suitable for geometry optimization of apo-CRABPII model structure. Similar procedures were employed for energy minimization and geometry optimization of slightly modified crystal structures of human glutathione *S*-transferase and the results were supported by enzymological studies (Hu *et al.*, 1997a,b).

Acknowledgments

We thank Drs A. Byrd, M. Jaskólski and A. Wlodawer for their critical reading of the manuscript. This work was supported by the National Cancer Institute, DHHS, under contract with ABL (X.C. and X.J.) and by NIH grant to H. Yan, Michigan State University (L.W., Y.L. and H.Y.). The contents of this publication do not necessarily reflect the views or policies of the Department of

Health and Human Services (DHHS) or the National Institute of Standards and Technology (NIST), nor does mention of trade names, commercial products, or organizations imply endorsement by DHHS, NIST, or the United States Government.

References

Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219–220.

Bailey, J. S. & Siu, C. H. (1988). Purification and partial characterization of a novel binding protein for retinoic acid from neonatal rat. *J. Biol. Chem.* **263**, 9326–9332.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. & Jones, T. A. (1994). Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. *Advan. Prot. Chem.* **45**, 89–151.

Bhat, T. N. (1988). Calculation of an OMIT map. *J. Appl. Crystallog.* **21**, 279–281.

Boylan, J. F. & Gudas, L. J. (1992). The level of CRABP-I expression influences the amounts and types of all-*trans*-retinoic acid metabolites in F9 teratocarcinoma stem cells. *J. Biol. Chem.* **267**, 21486–21491.

Brünger, A. T. (1992). *In X-PLOR (Version 3. 1) Manual,* pp. 187–218, Yale University Press, New Haven, CT.

Brünger, A. T. (1993). Assessment of phase accuracy by cross validation: the free *R* value. Method and applications. *Acta Crystallog. sect. D,* **49**, 24–36.

Bucco, R. A., Zheng, W. L., Davis, J. T., Sierra-Rivera, E., Osteen, K. G., Chaudhary, A. K. & Ong, D. E. (1997). Cellular retinoic acid-binding protein(II) presence in rat uterine epithelial cells correlates with their synthesis of retinoic acid. *Biochemistry,* **36**, 4009–4014.

Carson, M. (1987). Ribbon models of macromolecules. *J. Mol. Graph.* **5**, 103–106.

Elias, P. M. & Williams, M. L. (1981). Retinoids, cancer, and the skin. *Arch. Dermatol.* **117**, 160–168.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A,* **47**, 392–400.

Fiorella, P. D. & Napoli, J. L. (1991). Expression of cellular retinoic acid binding protein (CRABP) in *Escherichia coli*. Characterization and evidence that holo-CRABP is a substrate in retinoic acid metabolism. *J. Biol. Chem.* **266**, 16572–16579.

Fiorella, P. D. & Napoli, J. L. (1994). Microsomal retinoic acid metabolism. Effects of cellular retinoic acid-binding protein (type I) and C18–hydroxylation as an initial step. *J. Biol. Chem.* **269**, 10538–10544.

Furey, W., Wang, B. & Sax, M. (1982). Crystallographic computing on an array processor. *J. Appl. Crystallog.* **15**, 160–166.

Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. (1987). Identification of a receptor for the morphogen retinoic acid. *Nature,* **330**, 624–629.

Gorry, P., Lufkin, T., Dierich, A., Rochette-Egly, C., Decimo, D., Dolle, P., Mark, M., Durand, B. & Chambon, P. (1994). The cellular retinoic acid binding protein I is dispensable. *Proc. Natl Acad. Sci. USA,* **91**, 9032–9036.

Gudas, L. J., Sporn, J. & Roberts, A. B. (1994). Cellular biology and biochemistry of retinoids. In *THE RETINOIDS: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), 2nd edit., pp. 443–520, Raven Press, New York.

Haunerland, N. H., Jacobson, B. L., Wesenberg, G., Rayment, I. & Holden, H. M. (1994). Three-dimensional structure of the muscle fatty-acid-binding protein isolated from the desert locust *Schistocerca gregaria*. *Biochemistry,* **33**, 12378–12385.

Hendrickson, W. A. (1985). Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol.* **115**, 252–270.

Hodsdon, M. E. & Cistola, D. P. (1997a). Discrete backbone disorder in the nuclear magnetic resonance structure of apo intestinal fatty acid-binding protein: Implications for the mechanism of ligand entry. *Biochemistry,* **36**, 1450–1460.

Hodsdon, M. E. & Cistola, D. P. (1997b). Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N NMR relaxation and ^{1}H exchange. *Biochemistry,* **36**, 2278–2290.

Hong, W. K. & Itri, L. M. (1994). Retinoids and human cancer. In *THE RETINOIDS: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), 2nd edit., pp. 597–630, Raven Press, New York.

Howard, A. J., Gilliland, G. L., Finzel, B. Z., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Crystallog.* **20**, 383–387.

Hu, X., Ji, X., Srivastava, S. K., Xia, H., Awasthi, S., Nanduri, B., Awasthi, Y. C., Zimniak, P. & Singh, S. V. (1997a). Mechanism of differential catalytic efficiency of two polymorphic forms of human glutathione S-transferase P1-1 in the glutathione conjugation of carcinogenic diol epoxide of chrysene. *Arch. Biochem. Biophys.* **345**, 32–38.

Hu, X., O'Donnell, R., Srivastava, S. K., Xia, H., Zimniak, P., Nanduri, B., Bleicher, R. J., Awasthi, S., Awasthi, Y. C., Ji, X. & Singh, S. V. (1997b). Active site architecture of polymorphic forms of human glutathione S-transferase P1-1 accounts for their enantioselectivity and disparate activity in the glutathione conjugation of 7β,8α-dihydroxy-9α,10α-oxy-7,8,9,10-tetrahydrobenzo(*a*)pyrene. *Biochem. Biophys. Res. Commun.* **235**, 424–428.

Jamison, R. S., Newcomer, M. E. & Ong, D. E. (1994). Cellular retinoid-binding proteins: limited proteolysis reveals a conformational change upon ligand binding. *Biochemistry,* **33**, 2873–2879.

Jing, Y., Waxman, S. & Mira-y-Lopez, R. (1997). The cellular retinoic acid binding protein II is a positive regulator of retinoic acid signaling in breast cancer cells. *Cancer Res.* **57**, 1668–1672.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these methods. *Acta Crystallog. sect. A,* **47**, 110–119.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers,* **22**, 2577–2637.

Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid. *Structure,* **2**, 1241–1258.

Konnert, J. M. (1976). A restrained-parameter structure factor least-squares refinement procedure for large

asymmetric units. *Acta Crystallog. sect. A,* **32**, 614–617.
Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA,* **82**, 488–492.
Lampron, C., Rochette-Egly, C., Gorry, P., Dolle, P., Mark, M., Lufkin, T., LeMeur, M. & Chambon, P. (1995). Mice deficient in cellular retinoic acid binding protein II (CRABPII) or in both CRABPI and CRABPII are essentially normal. *Development,* **121**, 539–548.
Laskowski, R. A. (1993). PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
Lotan, R. (1980). Effects of vitamin A and its analogs (retinoids) on normal and neoplastic cells. *Biochim. Biophys. Acta,* **605**, 33–91.
Mangelsdorf, D. J., Ong, E. S., Dyck, J. A. & Evans, R. M. (1990). Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature,* **345**, 224–229.
Napoli, J. L., Boerman, M. H., Chai, X., Zhai, Y. & Fiorella, P. D. (1995). Enzymes and binding proteins affecting retinoic acid concentrations. *J. Steroid Biochem. Mol. Biol.* **53**, 497–502.
Navaza, J. (1994). An automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157–163.
Newcomer, M. E. (1995). Retinoid-binding proteins: structural determinants important for function. *FASEB J.* **9**, 229–239.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
Ong, D. E. & Chytil, F. (1975). Retinoic acid-binding protein in rat tissue. Partial purification and comparison to rat tissue retinol-binding protein. *J. Biol. Chem.* **250**, 6113–6117.
Ong, D. E., Newcomer, M. E. & Chytil, F. (1994). Cellular retinoid-binding proteins. In *THE RETINOIDS: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), 2nd edit., pp. 283–317, Raven Press, New York.
Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature,* **330**, 444–450.
Powell, M. J. D. (1977). Restart procedures for the conjugate gradient method. *Math. Prog.* **12**, 241–254.
Rizo, J., Liu, Z. P. & Gierasch, L. M. (1994). ^{1}H and ^{15}N resonance assignments and secondary structure of cellular retinoic acid-binding protein with and without bound ligand. *J. Biomol. NMR,* **4**, 741–760.
Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1989a). Crystal structure of rat intestinal fatty-acid-binding protein. Refinement and analysis of the *Escherichia coli*-derived protein with bound palmitate. *J. Mol. Biol.* **208**, 327–339.
Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1989b). Refined apoprotein structure of rat intestinal fatty acid binding protein produced in *Escherichia coli*. *Proc. Natl Acad. Sci. USA,* **86**, 7736–7740.
Sack, J. S. (1988). CHAIN: a crystallographic modeling program. *J. Mol. Graph.* **6**, 224–225.
Sani, B. P. & Hill, D. L. (1974). Retinoic acid: a binding protein in chick embryo metatarsal skin. *Biochem. Biophys. Res. Commun.* **61**, 1276–1282.
Scapin, G., Gordon, J. I. & Sacchettini, J. C. (1992). Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1. 2-Å resolution. *J. Biol. Chem.* **267**, 4253–4269.
Tallman, M. S. (1996). Differentiating therapy in acute myeloid leukemia. *Leukemia,* **10**(Suppl. 2), s33–38.
Thompson, J. R., Bratt, J. M. & Banaszak, L. J. (1995). Crystal structure of cellular retinoic acid binding protein I shows increased access to the binding cavity due to formation of an intermolecular β-sheet. *J. Mol. Biol.* **252**, 433–446.
Wang, L., Li, Y. & Yan, H. (1997). Structure-function relationships of cellular retinoic acid-binding proteins. Quantitative analysis of the ligand binding properties of the wild-type proteins and site-directed mutants. *J. Biol. Chem.* **272**, 1541–1547.
Winter, N., Bratt, J. M. & Banaszak, L. J. (1993). Crystal structures of holo and apo-cellular retinol-binding protein II. *J. Mol. Biol.* **230**, 1247–1259.
Xu, Z., Bernlohr, D. A. & Banaszak, L. J. (1992). Crystal structure of recombinant murine adipocyte lipid-binding protein. *Biochemistry,* **31**, 3484–3492.
Xu, Z., Bernlohr, D. A. & Banaszak, L. J. (1993). The adipocyte lipid-binding protein at 1. 6-Å resolution. Crystal structures of the apoprotein and with bound saturated and unsaturated fatty acids. *J. Biol. Chem.* **268**, 7874–7884.
Zhang, J., Liu, Z.-P., Jones, T. A., Gierasch, L. M. & Sambrook, J. F. (1992). Mutating the charged residues in the binding pocket of cellular retinoic acid-binding protein simultaneously reduced its binding affinity to retinoic acid and increases its thermostability. *Proteins: Struct. Funct. Genet.* **13**, 87–99.

Edited by I. A. Wilson

(Received 4 September 1997; received in revised form 11 February 1998; accepted 18 February 1998)