All-*trans* retinol, vitamin D and other hydrophobic compounds bind in the axial pore of the five-stranded coiled-coil domain of cartilage oligomeric matrix protein

Yuan Guo, Damir Bozic,
Vladimir N. Malashkevich¹,
Richard A. Kammerer, Therese Schulthess
and Jürgen Engel²

Abteilung für Biophysikalische Chemie, Biozentrum,
Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

¹Present address: Whitehead Institute for Biomedical Research,
Nine Cambridge Center, Cambridge, MA 02142, USA

²Corresponding author

The potential storage and delivery function of cartilage oligomeric matrix protein (COMP) for cell signaling molecules was explored by binding hydrophobic compounds to the recombinant five-stranded coiled-coil domain of COMP. Complex formation with benzene, cyclohexane, vitamin D₃ and elaidic acid was demonstrated through increases in denaturation temperatures of 2–10°C. For all-*trans* retinol and all-*trans* retinoic acid, an equilibrium dissociation constant $K_D = 0.6 \mu\text{M}$ was evaluated by fluorescence titration. Binding of benzene and all-*trans* retinol into the hydrophobic axial pore of the COMP coiled-coil domain was proven by the X-ray crystal structures of the corresponding complexes at 0.25 and 0.27 nm resolution, respectively. Benzene binds with its plane perpendicular to the pore axis. The binding site is between the two internal rings formed by Leu37 and Thr40 pointing into the pore of the COMP coiled-coil domain. The retinol β -ionone ring is positioned in a hydrophobic environment near Thr40, and the 1.1 nm long isoprene tail follows a completely hydrophobic region of the pore. Its terminal hydroxyl group complexes with a ring of the five side chains of Gln54. A mutant in which Gln54 is replaced by Ile binds all-*trans* retinol with affinity similar to the wild-type, demonstrating that hydrophobic interactions are predominant.

Keywords: fluorescence/hydrophobic pore/storage/X-ray crystallography

Introduction

A non-collagenous glycoprotein, cartilage oligomeric matrix protein (COMP), was first identified in cartilage (Hedbom *et al.*, 1992) but was later also found in tendon (Di Cesare *et al.*, 1994a; Smith *et al.*, 1997). The protein is a 524 kDa homopentamer of five subunits which consist of an N-terminal heptad repeat region (cc) followed by four epidermal growth factor (EGF)-like domains (EF), seven calcium-binding domains (T3) and a C-terminal globular domain (TC). According to this domain organization, COMP belongs to the family of thrombospondins. Heptad repeats (*abcdefg*)_n with preferentially hydrophobic

residues at positions *a* and *d* form α -helical coiled-coil domains (Cohen and Parry, 1994). Three subunits are oligomerized in this way in thrombospondins 1 and 2, but in thrombospondins 3 and 4 and in COMP a five-stranded coiled-coil is formed (Mörgelin *et al.*, 1992; Efimov *et al.*, 1994, 1995; Di Cesare *et al.*, 1994b; Hauser *et al.*, 1995; Qabar *et al.*, 1995).

Recently, the recombinant five-stranded coiled-coil domain of COMP (COMPcc) was crystallized and its structure was solved at 0.2 nm resolution (Efimov *et al.*, 1995; Malashkevich *et al.*, 1996). The most predominant feature of the coiled-coil is a 7.3 nm long hydrophobic pore with a diameter of 0.2–0.6 nm. Only two polar sites interrupt the hydrophobic lining of the pore. About 2.9 nm from the C-terminus, a ring of five glutamine side chains provides a putative ion trap to which a Cl[−] ion is bound and, ~1.5 nm from the N-terminus, a ring of OH groups of threonines side chains is located.

The function of COMP is not yet clear, but its importance is suggested by the recent discovery that mutations in the COMP gene are responsible for a genetic disorder in man, pseudoachondroplasia (Briggs *et al.*, 1995; Hecht *et al.*, 1995), with the phenotype being characterized by short stature and lax joints. Furthermore, increased COMP expression with increased weight bearing suggests an important role for COMP in providing strength in muscle and tendon (Smith *et al.*, 1997). In tissue sections of cartilage, COMP is localized to the territorial matrix of the chondrocytes, leading to speculation that it is involved in interactions with the cells (Hedbom *et al.*, 1992). Its bouquet-like structure with five terminal TC domains may provide the multivalency necessary for strong interactions with cell surface receptors or other proteins. COMP is revealed in the growth regions of cartilage, primarily in the proliferative zones, but is absent in the resting and hypertrophic zones (Ekman and Heinegard, 1992; Shen *et al.*, 1995). It was suggested, therefore, that COMP may be important for cellular proliferation. In addition, it was proposed that COMP is involved in signaling between external load and cellular response (Banes *et al.*, 1995; Smith *et al.*, 1997).

Vitamin D₃ and the morphogen all-*trans* retinoic acid have been shown to signal important events during morphogenesis and repair of cartilage and bone. Retinoic acid has the capacity to activate homeobox gene clusters (Mavilio *et al.*, 1988; Marshall *et al.*, 1996), to influence limb formation (Thaller and Eichele, 1987; Means and Gudas, 1995) and to stimulate matrix calcification and collagen synthesis in weight-bearing growth plate chondrocytes (Wu, 1997). The hormone 1,25-dihydroxyvitamin D₃ promotes cartilage and bone mineralization and is involved in the regulation of several proteins of these tissues. Furthermore, the nuclear vitamin D receptor is coupled functionally with the retinoid X receptor (RXR)

Fig. 1. Chemical equations of the hydrophobic compound used for binding to COMPcc. The C-atoms in retinol are numbered (Cowan *et al.*, 1990). The approximate length of the isoprene chain of all-*trans* retinol and all-*trans* retinoic acid is 1.1 nm, and the hydrocarbon chain of elaidic acid is 2.1 nm.

by heterodimer formation (Haussler *et al.*, 1997). 1,25-dihydroxyvitamin D₃ and retinoic acid are essentially insoluble in aqueous media and, for their transport and storage, binding proteins with hydrophobic interactions as well as specificity-determining interactions between ligand and protein have been described (Cowan *et al.*, 1993; Ray, 1996). The hydrophobicity and geometry of the channel in the coiled-coil domain of COMP suggest a binding potential for retinoids, vitamin D₃ and other hydrophobic compounds. This property appeared to be worth exploring in view of the many indications that COMP may be involved in signaling functions in which retinoids and vitamin D₃ may be involved. In the present work, binding of several hydrophobic compounds (Figure 1) was explored by employing changes of thermal stability of COMPcc, fluorescence titrations and X-ray crystallography.

Results

Binding of hydrophobic compounds increases the thermal stability of COMPcc

As shown in Figure 2 and Table I, small hydrophobic molecules such as benzene and cyclohexane increase the midpoint transition temperature T_m of unfolding of reduced and alkylated COMPcc by 2°C. For larger molecules such as vitamin D₃ and elaidic acid, this effect is more pronounced, and the change in T_m is 8 and 10°C, respectively. Oxidized COMPcc is more stable than reduced and alkylated COMPcc due to the disulfide ring connecting the five chains (Efimov *et al.*, 1995; Malashkevich *et al.*, 1996). The T_m is >100°C under non-denaturing conditions but is decreased to 60°C in the presence of 4.5 M guanidine-HCl. Measurements of the stabilization of oxidized COMPcc by hydrophobic molecules were therefore performed in the presence of this denaturant (Table I). Changes in T_m upon binding of vitamin D₃ and elaidic acid were somewhat smaller than for reduced and alkylated COMPcc.

Determination of binding affinities by fluorescence titration

Binding of all-*trans* retinol to oxidized COMPcc was measured by fluorescence titration using the fluorescence change of retinol. Measurements were performed with

Fig. 2. Increased thermal stability of the coiled-coil domain by saturation with hydrophobic compounds. Thermal denaturation profiles were recorded by circular dichroism at 221 nm for COMPcc in the reduced and alkylated state. The protein concentration was 2.69 μM in PBS (pH 7.5). Symbols denote COMPcc (○), COMPcc complexes with benzene (▲) and COMPcc complexes with elaidic acid (□).

Table I. Midpoint temperature T_m of the thermal unfolding of COMPcc in the absence of and after saturation with hydrophobic ligands

Protein	Ligand	T_m (°C)	ΔT_m (°C)
Reduced and alkylated COMPcc in PBS (pH 7.5)	–	48	0
	benzene	50	2
	cyclohexane	50	2
	vitamin D ₃	56	8
	elaidic acid	58	10
Oxidized COMPcc in PBS (pH 7.5) in the presence of 4.5 M guanidine-HCl	–	60	0
	vitamin D ₃	64	5
	elaidic acid	65	6

ΔT_m = increase in midpoint temperature caused by the addition of the ligand.

oxidized COMPcc and the designed protein in which Gln54 has been replaced by Ile (COMPccGln54Ile) (Figure 3A). The water-insoluble all-*trans* retinol was dissolved in dimethylsulfoxide (DMSO) and aliquots of this solution were added to the solution of the protein in phosphate-buffered saline (PBS). The final concentration of DMSO was <1% (v/v) and no precipitation of protein was observed. By a least square fit of a 1:1 binding mechanism (see Materials and methods) to the data in Figure 3A, dissociation equilibrium constants of $K_D = 0.6$ and 0.8 μM were derived for COMPcc and COMPccGln54Ile, respectively.

All-*trans* retinoic acid is not fluorescent under the conditions described above and its binding to COMPcc was determined by competition with all-*trans* retinol. Increasing quantities of retinoic acid in DMSO and a constant quantity of all-*trans* retinol in the same solvent were added to the aqueous protein solution. A concentration-dependent decrease in fluorescence intensity of the

Fig. 3. Equilibrium dissociation constants from fluorescence titration. (A) All-*trans* retinol was added to oxidized COMPcc (○) or its mutant COMPccGln54Ile (△) and the fluorescence was recorded at 480 nm (excitation at 330 nm). Protein concentrations were 0.18 μM in PBS (pH 7.5). The average K_D value which was evaluated for COMPcc is 0.6 μM, and for COMPccGln54Ile is 0.8 μM. (B) For fluorescence competition of retinoic acid with retinol, retinoic acid and retinol mixtures in DMSO were transferred into a solution of oxidized COMPcc solution in PBS (pH 7.5). Protein concentration was 0.5 μM. Data were recorded in the same conditions as described above.

retinol was observed with increasing amounts of retinoic acid (Figure 3B). A 2.5-fold molar excess of retinoic acid was required to compete for 50% of the retinol-binding sites, indicating that retinoic acid has a slightly lower affinity than retinol for COMPcc. A similar result was obtained for binding of all-*trans* retinoic acid to the COMPccGln54Ile. Equilibrium values were reached in all experiments, as demonstrated by kinetic experiments, and it was also shown that the times of retinol and retinoic acid binding and dissociation are <1 min (data not shown).

Crystal structure of the complexes with benzene and all-*trans* retinol

The localization of benzene and retinol in the complex with COMPcc was determined by X-ray crystallography. Benzene was found to bind inside the COMPcc channel as indicated by a flat electron density peak between side chains of Leu37 and Thr40 from five different subunits (Figure 4). This density is somewhat diffused, suggesting that orientation of the benzene ring is not unique. Indeed, its almost orthonormal orientation relative to the pentamer axis can be due to the water-mediated interaction between the moderately polar benzene molecule and hydroxyl groups of Thr40. Orientation of the benzene ring around the axis is ambiguous, as the 6-fold symmetry of benzene is incompatible with the 5-fold symmetry of COMPcc. The appearance of four instead of the potential five water molecules may reflect deviation from the perfect 5-fold non-crystallographic symmetry which has already been discussed by Malashkevich *et al.* (1996). A water molecule is positioned close to the benzene ring, ~0.45 nm away from its center. Interactions between the phenylalanine ring and polar side chains in globular proteins were described before (Burley and Petsko, 1986). Benzene binding requires displacement of one water molecule bound near this position in the native structure (Malashkevich *et al.*, 1996). This may explain why the thermal stabilization effect of benzene is relatively small. Longer exposure of COMPcc crystals to benzene (see Materials and methods) seems to have resulted in diffusion of more benzene molecules inside the pore, but severe crystal damage impaired the diffraction quality of COMPcc crystals and did not allow proper crystallographic analysis. Similar experiments with cyclohexane also indicated the appearance of the difference density peaks in the COMPcc pore. More accurate interpretation of this density was not, however, possible due to the higher degeneracy of binding modes and the non-planar nature of cyclohexane.

In contrast to the slightly degenerated binding of benzene and cyclohexane, only one predominant binding conformer of all-*trans* retinol to COMPcc exists, as shown by the electron density in Figure 5A. Herein, we observe an axial binding of retinol with the β-ionone ring located in the cavity between the Thr40 and Leu44 side chain ring system and the isoprene tail pointing towards the C-terminus of the COMPcc pore, with the C5-C6-C7-C8 dihedral angle of 80° being close to the structurally predominant value of ±60° and 180° (Cowan *et al.*, 1990). Deviation of this angle towards the 'ideal' case of ±60° would result in clashes of the β-ionone moiety with the Leu44 side chains. The hydroxyl group of retinol is being hydrogen bonded by two of the five Gln54 residues (0.23 and 0.25 nm) which previously have been described to act as an ion trap for chloride ions (Malashkevich *et al.*, 1996). Except for these two polar interactions, the overall nature of binding is hydrophobic (Figure 5B). The C16, C17 and C18 methyl side chains of the β-ionone ring are in close van der Waals contact to the C^δ and C^ε side chain atoms of three of the five Leu44 ring system residues at distances of 0.35, 0.33 and 0.31 nm, respectively. The C19 and C20 methyl groups of the isoprene side chain are located in the hydrophobic cavities between the Leu44–Val47 and Val47–Leu51 ring systems. The distance between methyl groups and distal side chain atoms are

Fig. 4. Stereoview of the part of the crystal structure of the COMPcc–benzene complex near the benzene-binding site. (A) View along the pentamer axis from its N-terminus and (B) perpendicular to the pentamer axis. Only one possible orientation of benzene around the pentamer axis is shown. One of the water molecules (shown as red points) binds below the benzene ring ~ 0.45 nm away from its center. Benzene and water molecules were omitted from the phase calculations, and the electron density map (cyan) was calculated with coefficients $F_o - F_c$, $i\alpha_c$ and plotted at the 3σ level.

between 0.33 and 0.37 nm. The binding of retinol in an unambiguous conformation towards COMPcc can be explained by the breakdown of the exact molecular 5-fold symmetry of COMPcc which has already been described for the native structure (Malashkevich *et al.*, 1996). Herein, as a consequence of crystal packing, the Leu44 and Leu51 side chains already exhibit different conformeric states which might promote non-symmetrical binding of retinol to COMPcc, with further structural rearrangement of the side chains due to the lack of molecular symmetry in retinol.

Discussion

Functional features

The present work defines COMP as a protein which, in addition to its established structural functions in cartilage and tendon (Hedbom *et al.*, 1992; Smith *et al.*, 1997), binds the hormones retinol, retinoic acid and vitamin D shown to be important in development. The exact physiological role of this binding remains elusive, but indirect data point to an involvement of COMP in cellular signaling (Banes *et al.*, 1995; Smith *et al.*, 1997). The present data suggest a storage and delivery function for this protein for hydrophobic hormones which otherwise,

because of their limited solubility, could not be delivered to their cellular targets. The retinoids (vitamin A) are important during development of organs and tissues and in particular in limb growth. Gradient formation of the hormones and the regulated genes has been recognized to be important (Means and Gudas, 1995; Tabin, 1995). A number of genes including the Hox genes (Marshall *et al.*, 1996) supply cells with the essential positional and functional information required for their migration to their appropriate destination and for their proper morphological development. Retinoids can affect the expression of Hox genes in cell lines and embryonic tissues, and their teratogenic effects in development, including limb formation, have been demonstrated in numerous studies.

Furthermore, it was found that retinoids act as growth factors on many cells also in adult tissues. It is often not clear in these studies whether the physiological effects are due to retinol, retinoic acid or other frequently easily convertible forms of the retinoids but, in most cases, the all-*trans* forms with a linear arrangement of conjugated double bonds are most effective (Means and Gudas, 1995). $1\alpha,25$ -dihydroxy vitamin D_3 is one of the key hormones for calcium and bone homeostasis, but it also has a function in the regulation of cellular differentiation and proliferation (Walters, 1992; Ray, 1996). In addition to

Fig. 5. Crystal structure of the COMPcc–all-*trans* retinol complex. (A) The final $2F_o - F_c$ electron density stereoplot of the axial view on the COMPcc pore with bound all-*trans* retinol. The density was contoured at the 1σ level. (B) Van der Waals surface dot representation of retinol binding to COMPcc.

the receptors for retinoids and for vitamin D, a large number of binding and transport proteins have been identified for both classes of hydrophobic and essentially insoluble hormones (Blomhoff *et al.*, 1990; Bouillon *et al.*, 1995).

The cellular cytosolic retinoid-binding proteins CRABP and CRABPII are believed to mediate transport to the nuclear retinoid receptors RARs and RXRs. Other cytosolic binding proteins of the types CRBP and CRBPII are thought to be involved in the conversion of retinol to its aldehyde and to retinoic acid. Extracellular transport and storage is mediated by serum proteins, of which plasma retinol-binding protein (RBP) (Cogan *et al.*, 1976; Soprano and Blaner, 1994) and vitamin D-binding protein (Verboven *et al.*, 1995; Swamy *et al.*, 1997) have been studied.

Structural comparison of COMPcc with other lipophilic proteins

RBPs are members of two families of lipophilic proteins (Cowan *et al.*, 1993). These families show different three-dimensional structures but have some topological

similarities. Serum RBP, β -lactoglobulin, bilin-binding protein and odorant-binding protein reveal an eight-stranded antiparallel β -barrel that encapsulates the ligand (Cowan *et al.*, 1990). Members of the other family include cellular RBPs (CRBPs), CRABP and bovine P2 myelin, which transports fatty acids. The proteins are built up from 10 antiparallel β -strands that fold to form an orthogonal barrel containing the ligand (Cowan *et al.*, 1993).

The β -structures of the lipophilic proteins are different from the α -helical coiled-coil domain to which retinol binds in COMP. In this sense, COMP and, most likely, also its homologs thrombospondin 3 and 4 constitute a new class of lipophilic proteins. There are, however, a number of common features. In COMPcc, the all-*trans* retinol is entirely surrounded by the hydrophobic axial pore and can be contacted by surface solvent only at the N-terminal opening of the pore. Also in CRBP and RBP, the cavity has essentially no connection to the external solvent. The accessible surface of the retinol ligand after binding is only 1–3% of its value before binding (Cowan *et al.*, 1990, 1993). Furthermore, in COMPcc, CRBP and RBP, the retinol binds in its all-*trans* form in which the

isoprene chain is linear. This feature may be of critical importance for delivery of retinoids in the all-*trans* configuration in which they are physiologically active. A third common feature is the hydrogen bonding of the terminal alcohol group of retinol to the amino bond of a glutamine side chain (Gln108 in CRBP and Gln54 in COMPcc). According to the K_D values of retinol binding to COMPcc and to the glutamine-free mutant protein, the hydrogen bond does not contribute significantly to binding energy. In this case, most of the energy originates from the hydrophobic interaction. Vitamin D-binding protein is homologous to serum albumin, α -fetoprotein and afamin. Diffracting crystals of vitamin D-binding protein were obtained (Verboven *et al.*, 1995), but the structure of this protein has not been solved yet and structural data about its binding site are unknown.

Binding affinity of COMP and other retinol-binding proteins

The dissociation equilibrium constant for retinol binding to COMPcc is $K_D = 0.6 \mu\text{M}$, a value comparable with the $K_D = 0.2 \mu\text{M}$ of serum RBP. In blood serum, the total concentration of vitamin A was reported to be $2 \mu\text{M}$, with most of it bound to RBP (Goodman, 1984; Blomhoff *et al.*, 1990). For a storage function in the extracellular space, COMP should have a similar K_D to that of the extracellular transport protein RBP. It was also demonstrated that the rate constants of association and dissociation are high, as needed for fast exchange. The K_D values for the CRBPs are lower by two orders of magnitude than for the serum RBP (Ong *et al.*, 1994). As demonstrated, COMPcc has a similar affinity for retinol and retinoic acid and also binds other hydrophobic compounds. This tolerance is shared by cellular retinoid-binding proteins (Blomhoff *et al.*, 1990) and is a general property of extracellular lipophilic proteins including vitamin D-binding protein (Ray, 1996). For COMPcc, we studied a broad repertoire of hydrophobic compounds. Benzene, cyclohexane and a fatty acid bind, but affinities were quantitated only for retinol and retinoic acid. Of structural interest is the complex with benzene which suggests electrostatic interactions between a water molecule bound to polar groups of the protein and the delta (δ) π -electrons of the benzene ring. By a geometric analysis of high-resolution protein structures, Burley and Petsko (1986) demonstrated similar interactions between aromatic and polar side chain groups. The functional implications of an amino-aromatic hydrogen bond were investigated for CRBP (Jamison *et al.*, 1995). An increase in denaturation temperature was also observed for binding of benzene to a two-stranded coiled-coil structure (Gonzales *et al.*, 1996). In this case, however, binding was accompanied by a switch from two to three strands. This was not the case for the five-stranded COMPcc because of its wide axial pore.

Conclusions and future research

Our data clearly prove that all-*trans* retinoids bind to the coiled-coil domain of COMP. Structural data, binding affinities and exchange rates are compatible with a storage function for COMP which, in view of the importance of this morphogen for the development of limbs, is an attractive possibility. The qualitatively observed vitamin

D binding to the hydrophobic axial channel in COMPcc may also be of importance. At present, physiological data on the binding of these factors to COMP in cartilage tissues are still lacking. A first step in this direction is binding studies of tritium-labeled retinol to recombinantly expressed intact COMP. The radiolabeled all-*trans* retinol co-migrated with the 450 kDa protein on non-denaturing polyacrylamide gels. Further work is in progress to demonstrate that also in full-length COMP, retinol binding occurs to the coiled-coil domain.

Materials and methods

Reagents

The coiled-coil domain comprising residues Gly27–Gly72 of rat COMP was prepared by heterologous recombinant gene expression in *Escherichia coli*. The sequence of the wild-type COMPcc (27–72) is:

a d a d a d a d a d a d
GDLAPQMLRELQETNAALQDVRELLRQVKEITFLKNTVMCEACG

in which residues in position *a* and *d* of the heptad repeats are indicated. In the expressed fragment, Gly27 is replaced by Met. Formation and exchange of disulfide bonds were catalyzed by a mixture of reduced and oxidized glutathione as described by Efimov *et al.* (1995). The reduced and alkylated fragment was prepared by incubation with 10 mM dithiothreitol (DTT) (1 h, 37°C then 3 min, 95°C) and 25 mM *N*-ethylmaleimide (NEM) (1 h, 37°C). The designed protein mutant COMPccGln54Ile in which the glutamine residue in position 54 was exchanged by isoleucine was prepared according to Ho *et al.* (1989) using two sets of primers. Set 1 (46-N, 5'-CTGCATATGGACCTAGC-CCACAG-3'; and COMPI-54 3', 5'-GATCTCCTTGACGATCTGT-CGCAAGAGCTCTCTCAC-3'), set 2 (46-C, 5'-TGCGGATCCTTA-TCCGCAAGCGTCAC-3'; and COMPI-54 5', 5'-CTTGCGACAGAT-CGTCAAGGAGATCACCTTCCTG-3'). The PCR and DNA manipulations for cloning were performed according to standard protocols (Sambrook *et al.*, 1989). All constructs were confirmed by dideoxy DNA sequencing. The recombinant proteins were dialyzed against PBS buffer (20 mM sodium phosphate, 200 mM sodium chloride, pH 7.5) at 4°C and stored at -20°C. At room temperature, the wild-type and mutant protein were pentameric as judged by circular dichroism (CD) spectroscopy and analytical ultracentrifugation. Both proteins exhibited 80–90% α -helicity (Efimov *et al.*, 1994).

The mixture of protein with the insoluble hydrophobic compounds benzene, cyclohexane, vitamin D₃ and elaidic acid in PBS (pH 7.5) was shaken for >3 days at 4°C in order to facilitate equilibration of binding. The mixtures were clarified by centrifugation at 18000 r.p.m. for 15 min before use.

All-*trans* retinol and elaidic acid were from Sigma, vitamin D₃ and all-*trans* retinoic acid from Fluka and all other reagents from Merck.

Thermal unfolding

CD was recorded with a Cary 61 spectropolarimeter (Varian) using thermostated quartz cells of 1 mm path length. Melting curves were recorded by monitoring the molar ellipticity $[\Theta]$ at 221 nm as a function of temperature, applying a heating rate of 90°C/h. The incubation time with the hydrophobic ligands was chosen to be long enough to achieve equilibrium transition curves.

Fluorescence titration

Fluorescence was measured with a FP 777 fluorometer (Jasco) at an excitation wavelength of 330 nm and an emission wavelength of 480 nm. Aliquots of a solution of all-*trans* retinol in DMSO were added to 2 ml of a solution of oxidized COMPcc or COMPccGln54Ile in PBS (pH 7.5) at 20°C. After each addition, the solution was stirred gently and allowed to equilibrate in the dark for 2 min. Equilibrium dissociation constants (K_D s) were evaluated assuming binding of a single ligand molecule L to the protein P forming the complex PL with

$$K_D = [L] \times [P] / [PL] \quad (1)$$

by

$$F_b/F_s = ([L]/K_D) / (1 + [L]/K_D) \quad (2)$$

in which $[L]$ is the free ligand concentration, F_b is the fluorescence signal of bound ligand at a certain ligand concentration and F_s is the

Table II. Data collection statistics

Compounds	No. of reflections measured	No. of independent reflections	R_{merge}^a	Completeness ^b of data	Maximal resolution
COMPcc-retinol	20 313	5464	13.2%	86.0% (87.6%)	0.27 nm
COMPcc-benzene	23 243	7554	15.1%	99.1% (98.2%)	0.25 nm

^a $R_{\text{merge}} = \sum_j |I_j(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_j I_j(\text{hkl})$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity over j reflections.

^bCompleteness of the highest resolution shell is given in parentheses.

plateau signal at saturation. First estimations of K_D and F_s were obtained assuming $[L] = [L]_0$ the total ligand concentration and $F_b = \Delta F = F_{\text{obs}} - F_0$. F_{obs} and F_0 are the fluorescence with and without protein. An iterative method was applied to improve the value: (i) a new value of the free ligand concentration was calculated by $[L] = [L]_0 - \alpha[P]_0$, where $\alpha = F_b/F_s$ and $[P]_0$ is the total protein concentration; (ii) the contribution of the free ligand to the fluorescence signal (F_0) was calculated from the fluorescence of ligand without protein and $F_b = F_{\text{obs}} - F_0$; (iii) with the new values of $[L]$ and F_b , F_s and K_D were evaluated from Equation 2. The iteration was repeated until no significant change of K_D was observed.

For competition binding experiments, separate DMSO solutions were prepared containing a constant concentration of all-*trans* retinol at 0.8 mM and increasing amounts of all-*trans* retinoic acid at 0–6.8 mM. For each data point, 5 μ l of the mixture was added to 2 ml of oxidized COMPcc in PBS (pH 7.5) and incubated for 1 h in the dark with gentle stirring. The concentrations of all-*trans* retinol and retinoic acid were determined by transferring to ethanol using extinction coefficients of 46 000/cm/M at 325 nm and 45 000/cm/M at 350 nm, respectively (Cogan *et al.*, 1976).

Crystallization and soaking experiments

Oxidized COMPcc (0.4 mg/ml) in 20 mM sodium phosphate, 200 mM NaCl (pH 7.5) was incubated with solid all-*trans* retinol in the dark for >3 days at 4°C and the mixture was clarified by centrifugation at 18 000 r.p.m. for 15 min. The solution was concentrated to 16 mg/ml of COMPcc by ultrafiltration using a Vivaspin 15 ml device (Vivascience). Crystals were obtained using the hanging drop method at room temperature. The reservoir solution contained 7–8% PEG 8000, 50 mM sodium cacodylate (pH 6.0), 50 mM sodium acetate and 50 mM calcium acetate. The hanging drop was prepared by mixing 2 μ l of the reservoir solution with an equal volume of COMPcc-retinol complex solution. Crystals appeared in 1 day and reached their optimal size in 1 week.

Direct diffusion of benzene or cyclohexane into the native COMPcc crystals (Efimov *et al.*, 1995) in solution was not effective due to the very limited solubility of these compounds in the crystal stabilizing solution. The complex could be prepared by incubating the crystals mounted in a standard way in an X-ray capillary in the presence of saturated vapor of benzene (or cyclohexane) as described by Eriksson *et al.* (1992) for a lysozyme mutant. Crystal cracking occurred within hours in the presence of benzene, severely impairing the quality of the diffraction data. Therefore, several crystals were tested before acceptable data were collected.

For both complexes, the crystal size was $\sim 0.2 \times 0.3 \times 0.5$ mm³. Both crystal forms belong to the monoclinic space group P2₁, with unit cell parameters for the benzene complex of $a = 4.00$ nm, $b = 5.03$ nm, $c = 5.59$ nm, and $\beta = 104.3^\circ$, and for the retinol complex of $a = 3.847$ nm, $b = 4.947$ nm, $c = 5.498$ nm, and $\beta = 103.84^\circ$, the latter complex being isomorphous to the wild-type crystal form.

Data collection and evaluation

All data were collected on a MARresearch™ image plate mounted on an Elliot GX20 rotating anode generator at a wavelength of $\lambda = \text{CuK}\alpha = 0.15418$ nm at room temperature. The X-ray intensities were evaluated and scaled with the program package DENZO/SCALEPACK (Otwinowski, 1993); data reduction was performed within the CCP4 package (Collaborative Computational Project, number 4, 1994). A summary of data collection statistics is given in Table II.

Model building and crystallographic refinement

The structure of the COMPcc-benzene complex was determined by the difference Fourier technique and refined to convergence using positional and individual B -factor refinement with the TNT routine (Tronrud *et al.*, 1987).

As the COMPcc-retinol complex crystallized completely isomorphous

Table III. Final refinement statistics of the COMPcc complexes

	COMPcc–retinol	COMPcc–benzene		
Mean standard deviation from ideality				
Bonds	0.0006 nm	0.0012 nm		
Angles	1.200°	1.480°		
Dihedrals	18.68	24.5		
Impropers	1.1	3.4		
Final <i>R</i> -factor	19.5%	18.8%		
Temperature factor distribution of the protein	No. of atoms	Average <i>B</i>	No. of atoms	Average <i>B</i>
Main chain	920	28.7	920	35.5
Side chain	905	31.3	905	45.4
All atoms	1825	29.9	1825	40.7
Ligands	21	28.8	6	37.4
Water	152	28.2	173	46.3
All atoms	1998	29.8	2010	41.3

to the wild-type structure, the dataset was submitted to a simulated annealing run with X-PLOR (Brünger, 1992). The resulting density clearly showed one additional molecule of retinol. The coordinates and crystallographic parameters for retinol were taken from the X-PLOR heteroatom dictionary and built into the density using the program MAIN (Turk, 1988). The model was submitted to energy-restrained crystallographic refinement with X-PLOR (Brünger, 1992) using a grouped B -factor refinement procedure. The protein parameters were taken from Engh and Huber (1991). The final crystallographic R -factor of the COMPcc-retinol complex is 19.46% using all data between 3.0 and 0.27 nm resolution. The quality of the final density is shown in Figure 5A. The stereochemical quality of the structure was assessed with the program PROCHECK (Laskowski *et al.*, 1993). Within the Ramachandran plot (Ramachandran and Sasiekharan, 1968), 94% (202 residues) are located in the most favored region, 4.6% (10 residues) are in the additionally allowed, 1.4% (three residues) in the generously allowed and none in the disallowed region.

A summary of the refinement statistics is given in the Table III. Figure 4 was drawn according to Jones *et al.* (1991), and Figure 5A and B was drawn with the program MAIN (Turk, 1988).

Co-ordinates are being submitted to the Brookhaven Protein Data Base and meanwhile are available upon request from the authors (engel@ubaclu.unibas.ch).

Acknowledgements

We thank Dr Patrik Maurer and Jochen Thur, Institut für Biochemie, Universität Köln for providing recombinant full-length rat COMP. This work was supported by grants from EU program Biotechnology II to J.E.

References

- Banes, A.J., Tsuzaki, M., Yamamoto, J., Fischer, T., Brigman, B., Brown, T. and Miller, L. (1995) Mechanoreception at the cellular level: the detection, interpretation and diversity of responses to mechanical stimuli. *Biochem. Cell Biol.*, **73**, 349–365.
- Blomhoff, R., Green, M.H., Berg, T. and Norum, K.R. (1990) Transport and storage of vitamin A. *Science*, **250**, 399–404.
- Bouillon, R., Okamura, W.H. and Norman, A.W. (1995) Structure–function relationships in the vitamin D endocrine system. *Endocr. Rev.*, **16**, 200–257.

- Briggs,M.D. *et al.* (1995) Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nature Genet.*, **10**, 330–336.
- Brünger,A.T. (1992) *X-PLOR, Version 3.1: A System for Crystallography and NMR*. Yale University Press, New Haven, CT.
- Burley,S.K. and Petsko,G.A. (1986) Amino–aromatic interactions in proteins. *FEBS Lett.*, **203**, 139–143.
- Cogan,U., Kopelman,M., Mokady,S. and Shinitzky,M. (1976) Binding affinities of retinol and related compounds to retinol binding proteins. *Eur. J. Biochem.*, **65**, 71–78.
- Cohen,C. and Parry,D.A.D. (1994) Alpha-helical coiled-coils: more facts and better predictions. *Science*, **263**, 488–489.
- Collaborative Computational Project, number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
- Cowan,S.W., Newcomer,M.E. and Jones,T.A. (1990) Crystallographic refinement of human serum retinol binding protein at 2 Å resolution. *Proteins*, **8**, 44–61.
- Cowan,S.W., Newcomer,M. E. and Jones,T. A. (1993) Crystallographic studies on a family of cellular lipophilic transport proteins. *J. Mol. Biol.*, **230**, 1225–1246.
- Di Cesare,P., Hauser,N., Lehman,D., Pasumarti,S. and Paulsson,M. (1994a) Cartilage oligomeric matrix protein (COMP) is an abundant component of tendon. *FEBS Lett.*, **354**, 237–240.
- Di Cesare,P., Mörgelin,M., Mann,K. and Paulsson,M. (1994b) Cartilage oligomeric matrix protein and thrombospondin 1. Purification from articular cartilage, electron microscopic structure and chondrocyte binding. *Eur. J. Biochem.*, **223**, 927–937.
- Efimov,V.P., Lustig,A. and Engel,J. (1994) The thrombospondin-like chains of cartilage oligomeric matrix protein are assembled by a five-stranded α -helical bundle between residues 20 and 83. *FEBS Lett.*, **341**, 54–58.
- Efimov,V.P., Engel,J. and Malashkevich,V.N. (1995) Crystallization and preliminary crystallographic study of the pentamerizing domain from cartilage oligomeric protein: a five-stranded α -helical bundle. *Proteins*, **24**, 259–262.
- Ekman,S. and Heinegard,D. (1992) Immunohistochemical localization of matrix proteins in the femoral joint cartilage of growing commercial pigs. *Vet. Pathol.*, **29**, 514–520.
- Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.*, **A47**, 392–400.
- Eriksson,A.E., Baase,W.A., Wozniak,J.A. and Matthews,B.W. (1992) A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature*, **355**, 371–373.
- Goodman,D.S. (1984) Plasma retinol binding protein. In Sporn,M.B., Roberts,A.B. and Goodman,D.S. (eds), *The Retinoids*. 1st edn. Academic Press, NY, Vol. 2, pp. 41–88.
- Gonzales,L., Jr, Plecs,J.J. and Alber,T. (1996) An engineered allosteric switch in leucine-zipper oligomerization. *Nature Struct. Biol.*, **3**, 510–515.
- Hauser,N., Paulsson,M., Kale,A.A. and Di Cesare,P.E. (1995) Tendon extracellular matrix contains pentameric thrombospondin-4 (TSP-4). *FEBS Lett.*, **368**, 307–310.
- Haussler,M.R., Haussler,C.A., Jurutka,P.W., Thompson,P.D., Hsieh,J.C., Remus,L.S., Selznick,S.H. and Whitfield,G.K. (1997) The vitamin D hormone and its nuclear receptor: molecular actions and disease states. *J. Endocrinol.*, **154**, 57–73.
- Hecht,J.T. *et al.* (1995) Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nature Genet.*, **10**, 325–329.
- Hedbom,E. *et al.* (1992) Cartilage matrix proteins. An acidic oligomeric protein (COMP) detected only in cartilage. *J. Biol. Chem.*, **267**, 6132–6136.
- Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
- Jamison,R.S., Kakkad,B., Ebert,D.H., Newcomer,M.E. and Ong,D.E. (1995) Test of the contribution of an amino–aromatic hydrogen bond to protein function. *Biochemistry*, **34**, 11128–11132.
- Jones,T.A., Zou,J.-W., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **D47**, 110–119.
- Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
- Malashkevich,V.N., Kammerer,R.A., Efimov,V., Schulthess,T. and Engel,J. (1996) The crystal structure of a five-stranded coiled-coil in COMP: a prototype ion channel? *Science*, **274**, 761–765.
- Marshall,H., Morrison,A., Studer,M., Poperl,H. and Krumlauf,R. (1996) Retinoids and Hox genes. *FASEB J.*, **10**, 969–978.
- Mavilio,F., Simeone,A., Boncinelli,E. and Andrews,P.W. (1988) Activation of four homeobox gene clusters in human embryonic carcinoma cells induced to differentiate by retinoic acid. *Differentiation*, **37**, 73–79.
- Means,A.L. and Gudas,L.J. (1995) The roles of retinoids in vertebrate development. *Annu. Rev. Biochem.*, **64**, 201–233.
- Mörgelin,M., Heinegard,D., Engel,J. and Paulsson,M. (1992) Electron microscopy of native COMP (cartilage oligomeric matrix protein) purified from the swam rat chondrosarcoma reveals a five-armed structure. *J. Biol. Chem.*, **267**, 6137–6141.
- Ong,D.E., Newcomer,M.E. and Chytil,F. (1994) Cellular retinoid-binding proteins. In Sporn,M.B., Roberts,A.B. and Goodman,D.S. (eds), *The Retinoids*. 2nd edn. Academic Press, NY, pp. 283–317.
- Otwinowski,Z. (1993) Data collection and processing. In Sawyer,L., Isaacs,N. and Bailey,S. (eds), *Proceedings of the CCP4 Study Weekend*. SERC Daresbury Laboratory, Warrington, UK, pp. 56–62.
- Qabar,A., Derick,L., Lawler,J. and Dixit,V. (1995) Thrombospondin 3 is a pentameric molecule held together by interchain disulfide linkage involving two cysteine residues. *J. Biol. Chem.*, **270**, 12725–12729.
- Ramachandran,G.N. and Sasiekharan,V. (1968) Conformation of polypeptides and protein. *Adv. Protein Chem.*, **23**, 283–487.
- Ray,R. (1996) Molecular recognition in vitamin D-binding protein. *Proc. Soc. Exp. Biol. Med.*, **212**, 305–312.
- Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Shen,Z., Heinegard,D. and Sommarin,Y. (1995) Distribution and expression of cartilage oligomeric matrix protein and bone sialoprotein show marked changes during femoral head development. *Matrix Biol.*, **14**, 773–781.
- Smith,R.K.W., Zunino,L., Webbon,P.M. and Heinegard,D. (1997) The distribution of cartilage oligomeric matrix protein (COMP) in tendon and its variation with tendon site, age and load. *Matrix Biol.*, **16**, 255–271.
- Soprano,D.R. and Blaner,W.S. (1994) Plasma retinol-binding protein. In Sporn,M.B., Roberts,A.B. and Goodman,D.S. (eds), *The Retinoids*. 2nd edn. Academic Press, NY, pp. 257–282.
- Swamy,N., Dutta,A. and Ray,R. (1997) Roles of the structure and orientation of ligands and ligand mimics inside the ligand-binding pocket of the vitamin D-binding protein. *Biochemistry*, **36**, 7432–7436.
- Tabin,C. (1995) The initiation of the limb bud: growth factors, Hox genes and retinoids. *Cell*, **80**, 671–674.
- Thaller,C. and Eichele,G. (1987) Identification and spatial distribution of retinoids in the developing chick limb bud. *Nature*, **327**, 625–628.
- Tronrud,D.E., Ten Eyck,L.F. and Matthews,B.W. (1987) An efficient general-purpose least squares refinement program for macromolecular structures. *Acta Crystallogr.*, **A43**, 489–501.
- Turk,D. (1988) Development and usage of a molecular graphics program. PhD thesis, University of Ljubljana, Slovenia.
- Verboven,C.C., De Bondt,H.L., De Ranter,C., Bouillon,R. and Van Baelen,H. (1995) Crystallization and X-ray investigation of vitamin D-binding protein from human serum. Identification of the crystal content. *J. Steroid Biochem. Mol. Biol.*, **54**, 11–14.
- Walters,M.R. (1992) Newly identified actions of the vitamin-D endocrine system. *Endocr. Rev.*, **13**, 719–764.
- Wu,L.N. (1997) Retinoic acid stimulates matrix calcification and initiates type I collagen synthesis in primary cultures of avian weight-bearing growth plate chondrocytes. *J. Cell Biochem.*, **65**, 209–230.

Received May 29, 1998; revised July 2, 1998;
accepted July 21, 1998