Pure & Appl. Chem., Vol. 70, No. 1, pp. 41–47, 1998.
Printed in Great Britain.

A structural view of ligand binding to the retinoid receptors

Bruno P. Klaholz and Dino Moras

Laboratoire de Biologie Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, BP 163, 67404 Illkirch Cedex, France.
Fax ++33 3 88 65 32 76. E-mail: moras@igbmc.u-strasbg.fr

Abstract: Nuclear receptors play an important role in transcription regulation. They bind as homo- or heterodimers to the response elements of their target genes and interact with numerous and diverse partners, *e.g.* coactivator and corepressor proteins, and transcription factors. Many of these processes are ligand-dependent, *i.e.* binding of natural ligands activates the nuclear receptor through conformational changes of the protein. Synthetic ligands can be made specific for a particular receptor and have the potential for reducing the side-effects of natural ligands in pharmaceutical applications. The crystal structures of ligand-binding domains of the retinoid receptors have brought the first insight into the spatial organization and the nature of the ligand-induced changes at the atomic level. Furthermore, these structures provide a starting point for structure-based drug design of retinoids.

The nuclear receptor superfamily

Cells use two principal pathways to transmit signals from the exterior through the cytoplasm to the nucleus. In the first one, membrane receptors bind ligands at the cell surface which do not cross the membrane themselves. For example, after the binding of growth factors to membrane receptors the signal is transmitted by a cascade involving serine- or tyrosine-kinases. The receptors for EGF (epidermal growth factor) and NGF (nerve growth factor) utilize water-soluble ligands such as peptide hormones and growth factors, where the mitogenic action of EGF is mediated by ligand-induced autophosphorylation of the EGF receptor. The polypeptidic neurotrophins promote differentiation and survival of peripheral and central neurons.

In the second pathway, the intracellular nuclear receptors control the activity of their target genes directly by binding to specific DNA-sequences called hormone response elements. Unlike membrane-bound receptors, nuclear receptors bind lipophilic hormones, which can cross the lipid bilayer of the cell membrane and then bind to their cognate receptors, either in the cytoplasm or the nucleus. These include receptors for steroid-hormones (ER, estrogen receptor, GR, glucocorticoid receptor, among others), vitamin D (VDR), thyroid hormone (TR) and retinoids (RAR and RXR), which are vitamin A metabolites. Not all these ligands are exclusively endocrine. Some of them need to be metabolically modified, as retinoic acid, others are completely synthesized in the cell, as prostaglandins. All these receptors are ligand-inducible transcription regulators, and are involved in development and cell differentiation (reviewed in refs. 1 and 2). Additionally, the so-called orphan receptors either are constitutively active or their ligands are not yet known.

It is interesting to note that the nuclear and membrane receptor pathways are connected, as for example retinoic acid (RA) and EGF (refs. 3 and 4), and RA and NGF (ref. 5), respectively, show cooperative effects on cell growth via RARs/RXRs.

Retinoid receptors

Nuclear receptors exhibit a modular structure with regions A through F, which correspond to distinct structural domains. The A-B regions contain the ligand-independent transactivation function 1 (AF-1).

Region C is the DNA-binding domain (DBD), comprising two highly conserved zinc fingers and the DBD dimerization surface. Region E is the ligand-binding domain (LBD), which harbours the main regulation functions, namely the ligand-binding site, the ligand-dependent activation function 2 (AF-2) and the LBD dimerization surface. The retinoid receptor DBD contains a 66-residue core whereas the size of the LBD is approximately 225 residues. These domains are linked by a short hinge region, which, by analogy to the steroid hormone receptors, may work as a nuclear translocation signal (refs. 6 and 7).

Although RARs and RXRs can form homodimers, the formation of heterodimers increases the affinity for the cognate response elements leading to the formation of anisotropic receptor-DNA-complexes (refs. 8 and 9). This polarity is due to the formation of dimers on direct repeats, whereas complexes on palindromes or everted repeats exhibit C_2-symmetry (ref. 10). Indeed, the core of the RAR and RXR response elements (RAREs and RXREs) is a direct repeat (DR) of the hexanucleotide PuGGTCA, spaced by 1, 2 or 5 nucleotides. With the exception of DR1, RXR occupies the 5'-position in RAR-heterodimers (DR2, DR5), even in complexes with VDR, TR or certain orphan receptors (refs. 1 and 11). In addition, these receptors interact with corepressors like SMRT (silencing mediator for RARs and TRs) (refs. 12 and 13) and N-CoR (nuclear receptor corepressor) (ref. 14), which can be released on ligand and DNA binding (reviewed in ref. 15). In order to link the nuclear receptors to the basal transcription machinery and thereby to the control of the transcription initiation complex and RNA-Polymerase II, coactivators are employed, *e.g.* SRC-1 (steroid receptor coactivator) (refs. 16 and 17) and TIF2 (transcriptional intermediary factor) (ref. 18). All these partners (proteins, DNA and ligands) lead to a multiplicity of interactions that end in an astonishingly specific action. As many of these processes are ligand-dependent the possibility to direct these interactions by synthetic ligands is an exciting perspective for pharmaceutical applications.

Metabolism of retinoids, and their significance in the organism

All RAR and RXR subtypes α, β and γ are activated by retinoic acid (RA), a vitamin A metabolite. The nutrient vitamin A (retinol) is transported in the serum by the retinol binding protein (RBP), and its cell uptake seems to be mediated by a putative membrane RBP-receptor. Then either a complex of retinol and the cellular RBP (CRBP-II) serves as a substrate for esterification by the lecithin:retinol acetyltransferase (for storage) or a CRBP-I complex is utilized by retinol dehydrogenase isozymes for the conversion to retinal, followed by the oxidation to RA by a cytosolic retinal dehydrogenase. RA can then be fixed by RA-binding proteins (CRABP), before being further catabolized (reviewed in refs. 7 and 19). It is interesting to note that on one hand the promoter of the CRABP-II gene contains both a RARE and a RXRE, and on the other hand the promoter of the RAR genes also have a RARE, leading to autoregulation of the retinoid signal. This complex interplay shows that retinoid receptors are part of a metabolite-controlled signalling system.

The use of retinoids in treatment of skin diseases including psoriasis, acne and skin cancer, in treatment of chronic myelogenous leukemia, cervical cancer and for cancer chemoprevention (refs. 7, 20 and 21) is related to the pleiotropic effects of RA on morphogenesis, differentiation and homeostasis, regulating key steps in cell growth, development and embryogenesis (refs. 22-24).

RARs bind both 9*cis*- and all-*trans* RA (9C-RA and AT-RA, respectively), whereas RXRs are selective for 9C-RA (refs. 25 and 26). As these two retinoid receptor response pathways activate different sets of genes, specific compounds can act selectively in these pathways and should at the same time reduce the side effects in therapeutical applications, emphasizing the need for structural information. The recent crystallographic investigation of the structures of the LBD of RXR (ref. 27) and RAR (ref. 28) in our laboratory has led to both an insight into the ligand-binding mechanism and the ligand specificity of these receptors, enabling a general sequence alignment of all known LBDs (ref. 29).

LBD structure, ligand binding and activation of retinoid receptors

The crystal structures of the LBDs of apo-hRXRα (ref. 27) and hRARγ bound to AT-RA (ref. 28) and 9C-RA (our unpublished results), have been determined at atomic resolution (2.7, 2.0 and 2.4Å, respectively). The first result of this studies was to show that both LBDs of RXR and RAR exhibit a novel common fold,

an antiparallel α-helical sandwich (Fig.1). The missing N- and C-terminal flanks are connected to the DBD and the F-domain, respectively. Helix H3, the β-turn, H4, H5, H6, H11 and H12 participate in the formation of the ligand-binding pocket. In the apo-RXR the C-terminal H12 and the Ω-loop between H2 and H3 point into the solvent, leading to a relatively loose structure sensitive to proteolysis, whereas the more compact holo-RAR LBD is significantly more stable (refs. 30 and 31). In the crystal structure the apo-RXR is seen as a dimer, thus providing information on the residues involved in the LBD dimerization surface (ref. 27), which is most probably conserved in heterodimers where RXR is a partner for RAR, VDR and TR.

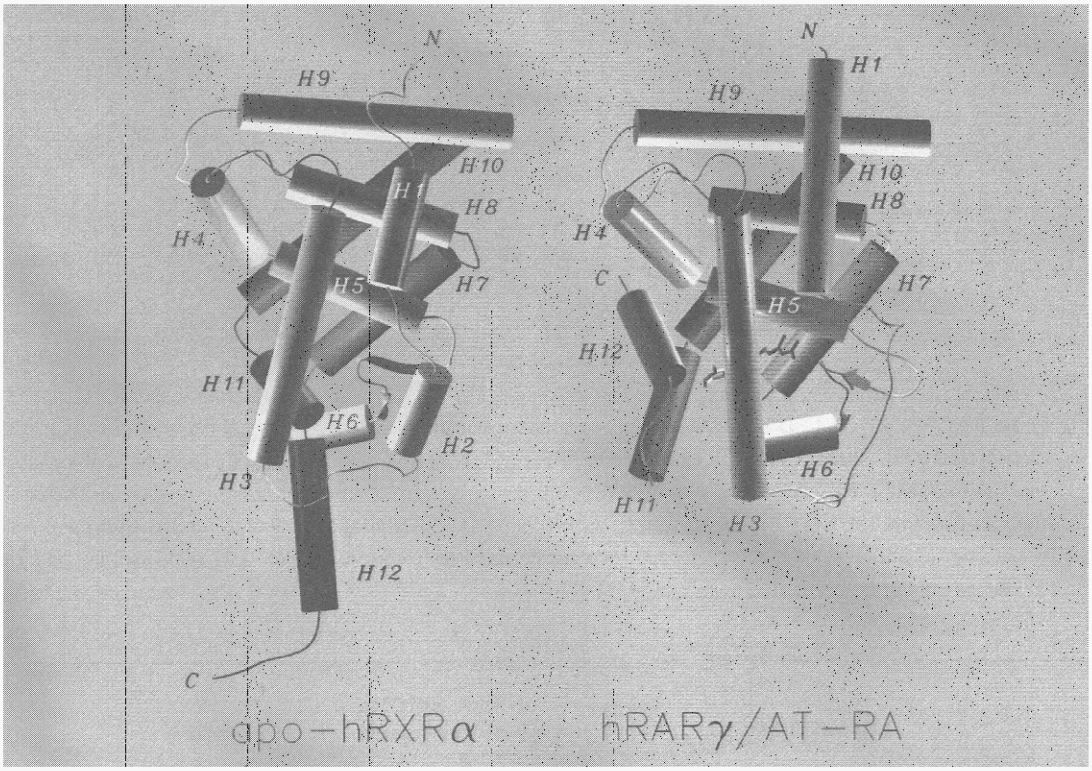

Fig.1 The ligand-binding domain of the retinoid receptors: comparison of the apo-hRXRα (ref. 27) with the holo-hRARγ bound to AT-RA (ref. 28), illustrating the common fold, the antiparallel α-helical sandwich (adapted from ref. 28). The numbering of the helices is indicated. All pictures were generated with the programme Setor (ref. 32).

The globally similar structures of the LBD of apo-RXR and holo-RAR suggest that the observed differences are due to ligand binding, although they share only 27% sequence homology. This ligand-induced conformational change is further confirmed by the above-mentioned protease mapping, apo-RAR-selective antibodies (ref. 34), and mutations in H12 (ref. 28), the helix containing the transactivation function AF-2. Compared to the apo-RXR, the Ω-loop undergoes a hinged-lid motion underneath the receptor in the holo-RAR, and is subsequently fixed by a salt bridge between Arg219 and Asp309 (H7). The amphipathic C-terminal helix H12 nestles against the receptor, being fixed by a strong salt bridge between Lys264 of H4 and Glu414, and to a lesser extent Glu417 (Fig.2A). The glutamates of the EMLE-motif, which form the core of the transactivation function AF-2, point to the surface of the receptor, whereas Met415 and Leu416 contact the ligand directly. The different views in Fig.2A and B illustrate how the retinoic acid is embedded between helices H3 and H5; note that the ligand is slightly bent and twisted by 43° between the β-ionone

moiety and the tetraene chain, which is due to the van der Waals contacts with the surrounding residues. The hydrophobic β-ionone moiety interacts with isoleucines, leucines, phenyalanines, alanines and methionines. The carboxylate group, however, is involved in a hydrophilic network, including salt bridges to Arg278 and Lys236, and hydrogen bonds to Ser289 (β-turn) and the main chain carbonyl group of Leu233 (H3) mediated by Lys236. Fig.3 schematically summarizes the residues interacting with the ligand.

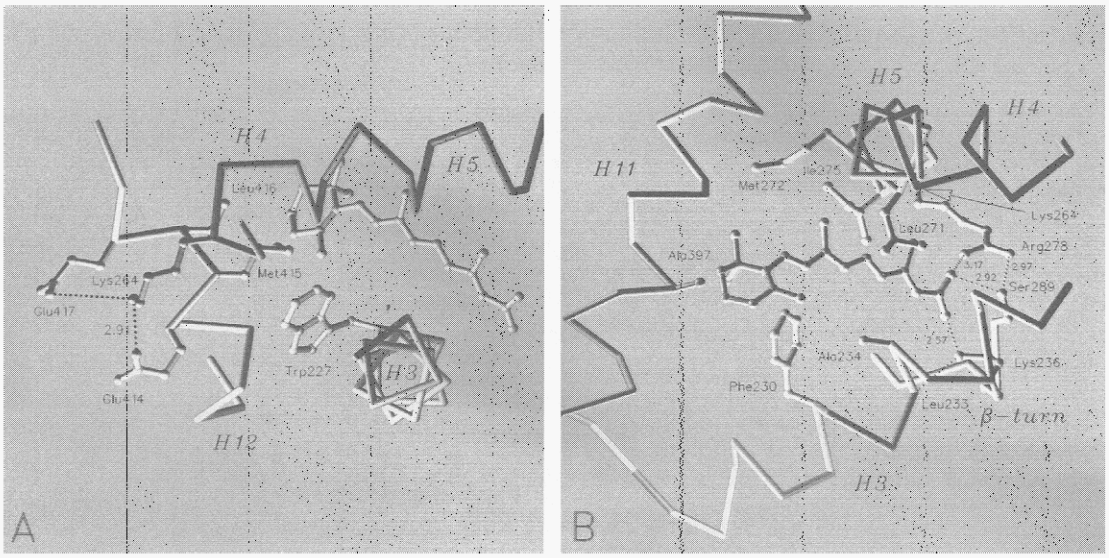

Fig.2 Detailed view of the LBD of hRARγ bound to all-*trans*-retinoic acid. **A** The EMLE-motif of the amphipathic transactivation helix H12, where Glu414 provides a strong salt bridge to Lys264 of helix H4, stabilizing the conformation of H12, and the methionine and leucine residues oriented towards the ligand. Trp227 has been used in tryptophan fluorescence quenching analysis (ref. 33). **B** AT-RA bound in the ligand-binding pocket, exhibiting hydrophobic contacts with the β-ionone moiety and a hydrophilic network of salt bridges and hydrogen bonds to the carboxylate group. H12 is omitted for clarity and would close the entry being salt bridged to the indicated Lys264.

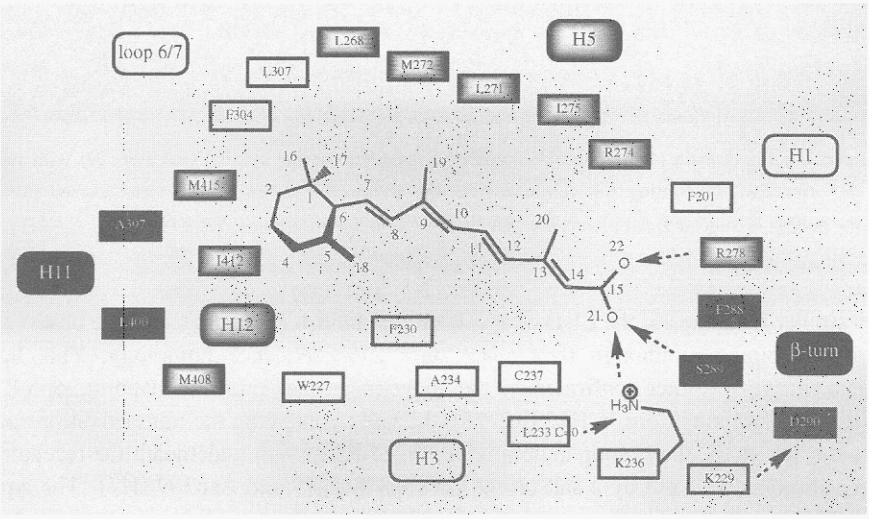

Fig.3 Schematic representation of the residues involved in hydrophobic contacts or salt bridges and H-bonds to the ligand (adapted from ref. 28).

Renaud et al. have proposed a mechanism of ligand entry based on an electrostatic field guidance (ref. 28). The negatively charged ligand appears to be attracted into the ligand-binding pocket by a positively charged cluster including Arg278, Lys236, Lys229, Lys240 and Arg274. Fig.4 shows the result of an electrostatic field calculation with the programme GRASP (ref. 35) where the ligand was omitted: the bundle of field lines leads to the charged cluster through the cavity, indicating the orientation of the ligand (added in the figure after the calculation). As Lys264 is located at the entrance of the cavity it could help to guide the ligand during its entry, before forming the salt bridge with Glu414. At the same time the receptor undergoes the conformational change mainly involving the refolding of the Ω-loop and H12, which closes the entry of the receptor after ligand binding. Dynamic simulations using molecular mechanics seem to confirm the proposed mechanism (A. Blondel et al., in preparation) but the mechanism may be different for uncharged ligands of other nuclear receptors.

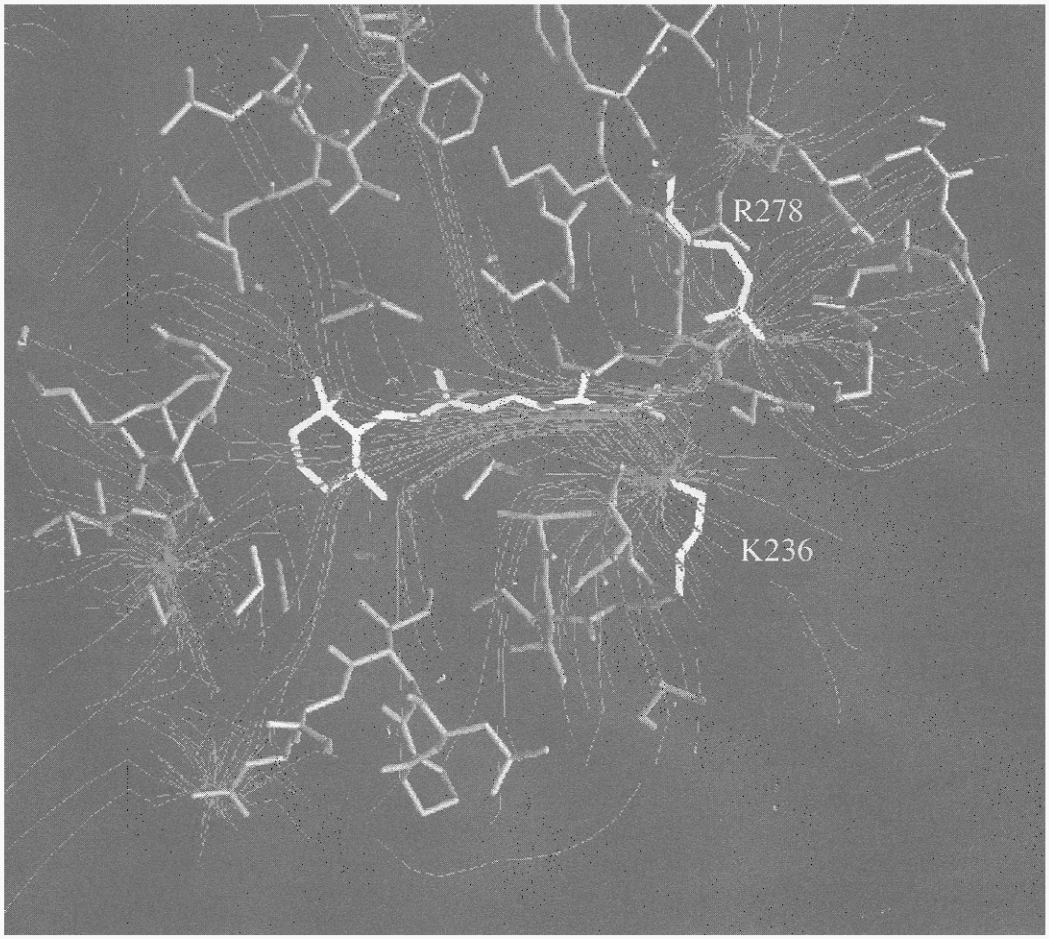

Fig.4 View of the ligand-binding pocket with the electrostatic field lines calculated with the programme GRASP (ref. 35) ending on the surface of the cavity and oriented from negative to positive potential. The ligand was omitted for the calculation and added for the figure (adapted from ref. 28).

The conformation of H12 in the holo-RAR is a characteristic of the active form of the receptor, first because only the holo-receptor is transcriptionally active, second the mutation K264A almost abolishes transactivation without impairing ligand and DNA binding, and third coactivators appear to interact only with the liganded receptor. This indicates that the conformational change generates a new surface enabling the formation of protein-protein complexes crucial for the transcription regulation, and also explains the ligand-induced release of corepressors.

On the basis of the common protein fold the present crystal structures have allowed a general sequence alignment of the whole nuclear receptor family. Thus a model of RXR and of GR could be proposed (ref. 29), and the homology modelling of other receptors is under way.

Retinoid receptors and ligand selectivity

Three types of selectivity have to be considered for nuclear receptor ligands: first they may be selective for families, *e.g.* steroid-hormone versus retinoid receptors, then for subfamilies, *e.g.* RAR versus RXR, and last but not least for subtypes, *i.e.* RARα, β and γ. Here we emphasize the second and third type.
Despite their different shapes, both retinoic acid isomers, AT-RA and 9C-RA, are apt to bind and transactivate RAR, whereas RXR binds exclusively 9C-RA. To understand how different ligands modulate the same receptor we have investigated the crystal structure of RAR LBD complexed with AT-RA (ref. 28) and 9C-RA (our unpublished results). The different shape of the natural isomers may hint at a more bent RXR ligand-binding pocket, as RXR shows no measureable affinity for the linear AT-RA, whereas it is selective for the more bent 9C-RA. Consistently, conformationally restricted RXR-selective compounds exhibit a pronounced bend when compared to RAR-selective compounds.

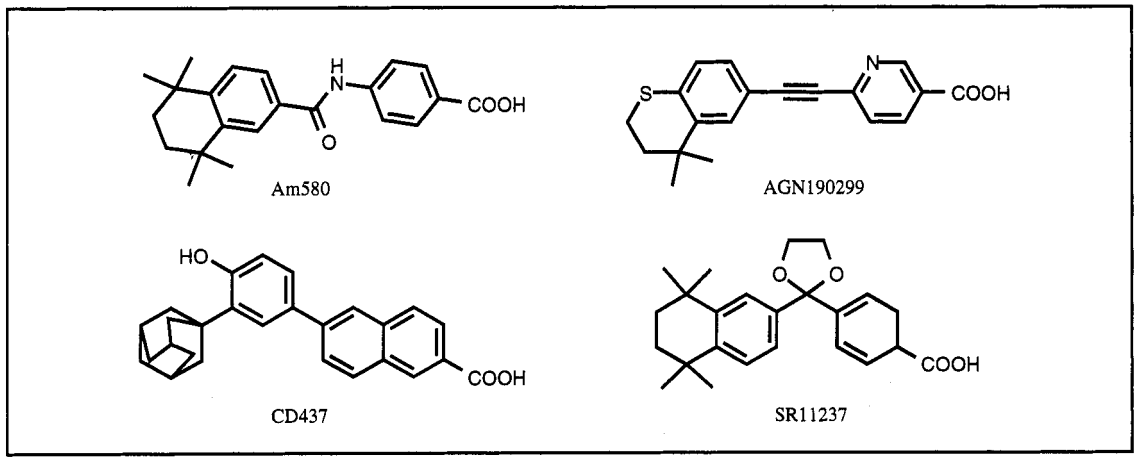

Fig.5 Some typical synthetic retinoids: the RARα selective Am580 (ref. 36), the RARβ/γ selective AGN190299 (ref. 37) and the RARγ selective CD437 (ref. 36), compared to the RXR-specific SR11237 (ref. 37).

Fig.5 shows four known synthetic retinoids, exhibiting RAR or RXR specificity and additionally RAR-subtype selectivity. The ligands contain on one hand a hydrophobic moiety, like the 5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl group, which mimicks the β-ionone moiety of RA, and on the other hand a carboxylate group which may be involved in a hydrophilic network similar to that of RA. The subtype selectivity of synthetic retinoids may be mainly due to residues in the ligand-binding pocket, as the sequence alignment of RAR-α, β and γ shows that only three residues differ in the ligand-binding pocket (ref. 28). In RARα the presence of Ser232, Ile270 and Val395 probably leads to a smaller cavity compared to RARβ (Ala225, Ile263, Val388) and RARγ (Ala234, Met272, Ala397) and may in part explain the RARα selectivity of the relatively small Am580 (our unpublished results). The reasons for retinoid binding and activation characteristics may be revealed by structural investigations of additional RAR- and RXR-ligand complexes.
Antagonists may function by steric clash of bulky substituents with residues of the ligand-binding pocket, especially with H12. This could disturb the position of H12 and hinder the formation of the active receptor conformation, thereby preventing the interaction with coactivators.
The question of selectivity stresses the need for further structural information, as the bound ligands could be used as templates for drug design. The present results provide a structural basis for the development of more selective RAR ligands, and will hopefully help to produce pharmaceutically relevant compounds.

Acknowledgments

We thank J. P. Renaud for kindly providing Fig.4, and J. G. Arnez and J. P. Renaud for careful reading of the manuscript. B. P. Klaholz is financially supported by the Deutscher Akademischer Austauschdienst, Doktoranden-stipendium aus Mitteln des dritten Hochschulsonderprogramms HSPIII.

References

1. Mangelsdorf, D. J., Thummel, C., Beato, M., et al. *Cell* **83**, 835-839 (1995).
2. Mangelsdorf, D. J. & Evans, R. M. *Cell* **83**, 841-850 (1995).
3. Zheng, Z. S., Polakowska, R., Johnson, A. & Goldsmith, L. A. *Cell Growth. Differ.* **3**, 225-232 (1992).
4. Roulier, S., Rochette-Egly, C., Alsat, E., Dufour, S., Porquet, D. & Evain-Brion, D. *Mol. Cell Endocrinol.* **118**, 125-135 (1996).
5. Plum, L. A. & Clagett-Dame, M. *Dev. Dyn.* **205**, 52-63 (1996).
6. Ylikomi, T., Bocquel, M. T., Berry, M., Gronemeyer, H. & Chambon, P. *EMBO J.* **11**, 3681-3694 (1992).
7. Giguere, V. *Endocr. Rev.* **15**, 61-79 (1994).
8. Zechel, C., Shen, X. Q., Chen, J. Y., Chen, Z. P., Chambon, P. & Gronemeyer, H. *EMBO J.* **13**, 1425-1433 (1994).
9. Kurokawa, R., Di Renzo, J., Boehm, M., et al. *Nature* **371**, 528-531 (1994).
10. Gronemeyer, H. & Moras, D. *Nature* **375**, 190-191 (1995).
11. Rastinejad, F., Perlmann, T., Evans, R. M. & Sigler, P. B. *Nature* **375**, 203-211 (1995).
12. Chen, J. D. & Evans, R. M. *Nature* **377**, 454-457 (1995).
13. Chen, J. D., Umesono, K. & Evans, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 7567-7571 (1996).
14. Hörlein, A. J., Naar, A. M., Heinzel, T., et al. *Nature* **377**, 397-404 (1995).
15. Horwitz, K. B., Jackson, T. A., Bain, D. L., Richer, J. K., Takimoto, G. S. & Tung, L. *Mol. Endocrinol.* **10**, 1167-1177 (1996).
16. Onate, S. A., Tsai, S. Y., Tsai, M. J. & O'Malley, B. W. *Science* **270**, 1354-1357 (1995).
17. McInerney, E. M., Tsai, M. J., O'Malley, B. W. & Katzenellenbogen, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10069-10073 (1996).
18. Voegel, J. J., Heine, M. J., Zechel, C., Chambon, P. & Gronemeyer, H. *EMBO J.* **15**, 3667-3675 (1996).
19. Napoli, J. L. *FASEB. J.* **10**, 993-1001 (1996).
20. Niles, R. M. *Adv. Exp. Med. Biol.* **375**, 1-15 (1995).
21. Tallman, M. S. & Wiernik, P. H. *J. Clin. Pharmacol.* **32**, 868-888 (1992).
22. Kastner, P., Mark, M. & Chambon, P. *Cell* **83**, 859-869 (1995).
23. De Luca, L. M. *FASEB. J.* **5**, 2924-2933 (1991).
24. Chambon, P. *FASEB. J.* **10**, 940-954 (1996).
25. Allenby, G., Bocquel, M. T., Saunders, M., et al. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 30-34 (1993).
26. Levin, A. A., Sturzenbecker, L. J., Kazmer, S., et al. *Nature* **355**, 359-361 (1992).
27. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. *Nature* **375**, 377-382 (1995).
28. Renaud, J. P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., Moras, D. *Nature* **378**, 681-689 (1995).
29. Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D., Gronemeyer, H. *Nat. Struct. Biol.* **3**, 87-94 (1996).
30. Leid, M. *J. Biol. Chem.* **269**, 14175-14181 (1994).
31. Keidel, S., Le Motte, P. & Apfel, C. *Mol. Cell Biol.* **14**, 287-298 (1994).
32. Evans, S. V. *J. Mol. Graphics* **11**, 134-138 (1993).
33. Lupisella, J. A., Driscoll, J. E., Metzler, W. J. & Reczek, P. R. *J. Biol. Chem.* **270**, 24884-24890 (1995).
34. Driscoll, J. E., Seachord, C. L., Lupisella, J. A., Darveau, R. P. & Reczek, P. R. *J. Biol. Chem.* **271**, 22969-22975 (1996).
35. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins* **11**, 281-286 (1991).
36. Bernard, B. A., Bernardon, J. M., Delescluse, C., et al. *Biochem. Biophys. Res. Commun.* **186**, 977-983 (1992).
37. Hembree, J. R., Agarwal, C., Beard, R. L., Chandraratna, R. A. & Eckert, R. *Cancer Res.* **56**, 1794-1799 (1996).