AD_____

Award Number: DAMD17-98-1-8133

TITLE: Toward a New Chemotherapy for Breast Cancer: Structural and Functional Mechanism of the Retinoid Receptors Addressed by a Novel Computer Approach

PRINCIPAL INVESTIGATOR: Ruben A. Abagyan, Ph.D.

CONTRACTING ORGANIZATION: The Scripps Research Institute
La Jolla, California 92037

REPORT DATE: May 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2000	Annual (1 May 99 - 30 Apr 00)

4. TITLE AND SUBTITLE
Toward a New Chemotherapy for Breast Cancer: Structural and Functional Mechanism of the Retinoid Receptors Addressed by a Novel Computer Approach

5. FUNDING NUMBERS
DAMD17-98-1-8133

6. AUTHOR(S)
Ruben A. Abagyan, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Scripps Research Institute
La Jolla, California 92037

E-MAIL:
ruben.abagyan@med.nyu.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Report contains color graphics.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The Retinoid Receptors have been recognized for many years as targets for breast cancer therapy both in estrogen receptor-positive and -negative cells. However, their clinical use is limited, so far, because of toxic side effects. A better understanding at the biochemical and structural level of the molecular mechanism of retinoid action could result in the development of molecules with improved potency and reduced toxicity. We have undertaken a variety of computer modeling strategies to analyze the structural determinants for receptor and isoform specificity and the structural mechanism of agonist activity. Our efforts have already resulted in the identification of novel RAR agonists presenting very original features and we initiated a collaboration with renowned chemists and biologists to build on our findings and develop lead molecules against breast cancer.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	52
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

N/A For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature 5-19-00 Date

Table of Contents

Introduction

Estrogen Receptor modulators, such as Tamoxifen, represent one of the most accepted therapy for the prevention and treatment of breast cancer. However, the resistance of estrogen receptor-negative tumor cells to treatment, and the possibility that Tamoxifen may increase the chances of endometrial carcinoma, create the need for improved therapeutic strategies. The Retinoid Receptors, two other members of the Nuclear Receptors (NR) family, each composed of three isoforms (α, β and γ), have been recognized for many years as targets for breast cancer therapy (Thacher et al for review). Retinoids exert their anticancer activities both in estrogen receptor-positive and -negative cells, mainly through their abilities to modulate the growth, differentiation, and apoptosis of tumor cells. Although retinoids are thought to have great therapeutic potential, their clinical use is so far limited, because of toxic side effects (Rizvi et al.). Consequently, the identification of novel structural motifs of small molecule modulators of the retinoid receptors is critical for the development of novel ligands with improved toxicity profiles, which could lead the way towards a novel generation of drugs against breast cancer. Our goal is to use state-of-the-art computer modeling technologies to develop such compounds.

Body

Task 1: Identify the structural determinants which confer a ligand retinoid X receptor (RXR) or retinoic acid receptor (RAR) selectivity

As detailed in the previous annual report, we were able to build a model of the RXR ligand binding domain (RXR-LBD) bound to 9-cis retinoic acid (9-cisRA), and docked RXR and RAR ligands to the two receptors. We predicted that the ligand binding pocket of RXR is shorter and bulkier than that of RAR, which accounts for receptor selectivity. Our prediction was recently confirmed by the crystal structure of the RXR/9cis RA complex (Gampe et al.).

As previously described, we also docked RXR- and RAR selective ligands into the receptors pocket and showed that our models correlated with experimental observations: non-binders had a positive calculated Van der Waals energy of interaction, whereas binders had a negative Van der Waals energy of interaction.

An increasing amount of data suggest that the RARα isoform, which controls the expression of RARβ, is the most relevant target for breast cancer therapy (Liu et al., Widschwendter et al.). A detailed understanding, at the atomic level, of the determinants for isoform specificity is therefore important to design ligands with appropriate receptor selectivity. We docked a RARα specific ligand, Am580, in the crystal structure of RAR γ, and in a model that we built of RARα and RARβ, and showed that residue 234, one of the three residues of the ligand binding pocket which are not conserved among the three RAR isoforms, was responsible for receptor isoform specificity (Figure 1).

Task 2: Identify the structural properties a ligand must satisfy to be an agonist

The structural mechanism of agonist and antagonist activity was recently revealed by a series of

crystal structures (Moras and Gronemeyer for review). Upon binding of an agonist, the C-terminal helix 12 of the LBD undergoes a conformational change, closes like a lid on the ligand, and forms a hydrophobic pocket on the surface of the receptor. This pocket is the target for several known and unknown coactivator proteins, which bind the LBD through a conserved LxxLL motif. In the antagonist-bound conformation, the helix H12 is pushed away from the ligand binding pocket, and occupies the coactivator binding site, preventing coactivator recruitment.

As described in our previous report, we collaborated with Dr Herbert Samuel's lab, at NYU Medical Center, to identify a novel coactivator, NRIF3, mediating functional specificity of nuclear hormone receptors. This novel protein coactivates RXR but not RAR. Our modeling suggests that NRIF3 binds to liganded receptors through an LxxIL module contained within the C-terminal domain of the coactivator. This data was the object of an article published in *Molecular and Cellular Biology* and is appended to this report (Article 1).

Very interesting data recently published on the structure-activity relationship of another nuclear receptor, the estrogen receptor, encouraged us to initiate an additional avenue of investigation on RAR, with possible important applications in breast cancer therapy. Paige et al. showed that different agonists of the estrogen receptor (estradiol, estriol, and diethylstilbestrol) could induce different conformational changes of the receptor LBD, resulting in recruitment of different coactivators (Paige et al.). Structural analysis of many nuclear receptors suggests that, upon ligand binding, the core of the receptor LBD remains fixed, while the C-terminal H12 helix is relocated and interacts with the ligand, thus generating a binding site for coactivator proteins (Moras and Gronemeyer, L. Freeman). It is therefore reasonable to predict that the different estrogen receptor agonists induce a different rearrangement of the H12 helix.

We built low energy models of the RAR LBD where the H12 helix is relocated onto different docking sites at the surface of the receptor. Only a limited number of rearrangements of H12 are possible (Figure 2). We plan to apply our virtual screening technology to these structures (see Task 3), in order to discover novel RAR agonists that could induce the recruitment of different coactivators. Such molecules could have a different pharmacological profile, compared with existing RAR ligands.

Task 3: Design rationally and test novel drugs against breast cancer

1-Use our model to screen libraries of compounds and select ligands with desired structural characteristics

As described in our previous report, we have built a model of antagonist-bound RAR and carried out a high-throughput computer screening where over 150,000 ligands were automatically docked into the receptor's binding pocket and assigned a score according to the quality of the fit. About thirty RAR antagonist candidates were selected for *in vitro* assay. This work was published in February in the Proceedings of the National Academy of Sciences and is appended to this report (Article 2).

We have also used a benchmark virtual library made of the structure of known ligands for

different members of the nuclear receptor family in order to optimize our virtual library screening technology. We could show that after four independent screenings, all the RAR ligands, and only the RAR ligands were selected (Table 1). This work was also published (Article 2).

Additionally, we have used the crystal structure of agonist-bound RAR to screen our virtual library of over 150,000 ligands and selected about 30 agonist candidates to be tested *in vitro*. This work was recently accepted for publication with revisions at the *Journal of Medicinal Chemistry*, and is appended to this report (Article 3).

Finally, we have used our model of agonist-bound RXR to screen our virtual compound library. However, we became aware during the screening that the crystal structure of the RXR-LBD/9cis retinoic acid complex was about to be published (R Gampe et al.). As a result, we postponed the *in vitro* testing of the RXR agonist candidates, and waited for the coordinates of the crystal structure of the receptor to carry out a novel screening which should result in an improved selection of ligands.

2- Test these ligands *in vitro* for their receptor selectivity, and their ability to trigger transactivation

As described in our previous report, two novel RAR antagonists were discovered, thus validating our modeling strategy (Article 2).

We also discovered novel agonists of RAR (Article 3). One of them, agonist 3, presents very original structural features:(i) a penta-methylated benzene linked by a ketone to a second aromatic ring, forms a very large hydrophobic head. (ii) A pyridine probably makes a strong interaction with Cys 237 and replaces a carboxylate or ester present in all RAR ligands described so far (Figure 3). These novel structural features may translate into improved toxicity profiles and result in the development of new ligands for breast cancer therapy.

This last finding is extremely encouraging, and we initiated a collaboration with Pr. Stephen Wilson and Pr. Herbert Samuels to discover analogs of agonist 3 with improved activity. Letters of intent from Drs. Wilson and Samuels are appended. Dr Wilson is Professor of chemistry at New York University and the director of the New York University Laboratory for Combinatorial Chemistry. He is a renown expert in solid phase chemistry, and the author of Combinatorial Chemistry: Synthesis and Application, John Wiley & Sons, March 1997; edited by Wilson & Czarnik, one of the first books on combinatorial chemistry and its organic chemistry applications. Pr. Samuels is an internationally recognized expert in endocrinology, and has been collaborating with us since the beginning of our research effort. Dr Matthieu Schapira, under Pr. Ruben Agabyan's supervision, coordinates this collaboration.

RAR ligands described in the literature belong to limited series of related structures, and display similar toxic side effects. The discovery of agonist 3 and the model of its interaction with RAR open an avenue towards new RAR agonist ligands with possibly more favorable specificity and toxicity profiles. We are anxious to build on our findings and, with the help of chemists and biologists, design, synthesize and test analogs of this novel structure with improved activity and specificity. This work could have direct applications in breast cancer therapy.

3- Test these ligands *in vivo* for their efficacy against rat mammary tumor

We plan to test *in vivo* the ligands discovered in collaboration with Dr. Herbert Samuels.

Key Research Accomplishments

-We built a model of the active conformation of RXR, which was recently confirmed by a crystal structure, and could derive the structural differences between the RAR and RXR binding pockets responsible for ligand selectivity (as illustrated in the previous annual report).

-We provided structural determinants for RAR isoforms selectivity (Figure 1).

-We modeled the interaction between RAR or RXR and NRIF3, a novel nuclear receptor coactivator (Article 1).

-We built predictive models of novel conformations of active RAR and plan to use these models to derive novel types of agonists (Figure 2).

-We discovered novel RAR antagonists (Article 2).

-We optimized our virtual screening technology and demonstrated its efficacy at discriminating between binders and non-binders (Table 1 and Article 2).

-We discovered an RAR agonist with very original structural features, and initiated a collaboration with chemists (Pr. S. Wilson's laboratory) and biologists (Pr. H. Samuels' laboratory) to design, synthesize and test structural analogs of this very promising molecule (Figure 3 and Article 3).

Reportable Outcomes

Article 1:
NRIF3 is a novel co-activator mediating functional specificity of nuclear hormone receptors
D. Li, V. Desai-Yajnik, E. Lo, M. Schapira, R.A. Abagyan, and H.H. Samuels
Molecular and Cellular Biology 1999 Oct;19(10):7191-7202

Article 2:
Rational discovery of novel nuclear receptor antagonists
M. Schapira, B. Raaka, H.H. Samuels, and R.A. Abagyan
Proceedings of the National Academy of Sciences U S A. 2000 Feb 1;97(3):1008-13

Article 3:
In Silico discovery of novel retinoic acid receptor agonist structures
M. Schapira, B. Raaka, H.H. Samuels, and R.A. Abagyan
Accepted for publication with revisions to *Journal of Medicinal Chemistry* April 2000

A patent filing application was sent to the Office of Industrial Liaison of New York University School of Medicine, to protect our discovery of Agonist 3.

Conclusions

During the first two years of our effort, we have accomplished most of the tasks outlined in the approved Statement of Work. We have identified the structural determinants which confer a ligand RXR or RAR selectivity, and successfully challenged our models with known ligands for either receptor. We have also put a rational on RAR isoforms specificity which will be useful to develop RARα specific agonists -the biological activity which seems the most promising for breast cancer therapy.

Based on recent structural data on the estrogen receptor, we also produced models of distinct RAR conformations, which we will use to derive novel types of ligands for this receptor. Such ligands could have interesting activities, mirroring the complexity of small molecule estrogen receptor modulators.

Finally, we have developed a strategy to rationally design RAR antagonists, and were also able to discover novel RAR agonists with very original structural features. Our current effort to develop analogs of these novel molecules could result in the identification of leads against breast cancer.

"So what":

We have developed a series of approaches to identify new antagonists and new agonists for the retinoid receptors, based on the structure of the receptor, and regardless of existing compounds. This original strategy already resulted in the discovery of active ligands presenting very original features. We are currently working in collaboration with chemists and biologists to develop analogs of these molecules with improved activity and specificity, which would constitute valuable leads for breast cancer therapy.

References

Freeman LP (1999) Increasing the Complexity of Coactivation in Nuclear Receptor Signaling, *Cell* **97**:5-8

Gampe R, V Montana, M Lambert, A Miller, R Bledsoe, M Milburn, S Kliewer, T Willson, and H.E Xu (2000), Asymmetry in the PPAR γ /RXRα Crystal Structure Reveals the Molecular Basis of Heterodimerization among Nuclear Receptors, *Molecular Cell*, **5**: 545-555,

Liu Y, Lee MO, Wang HG, Li Y, Hashimoto Y, Klaus M, Reed JC, Zhang X (1996) Retinoic acid receptor β mediates the growth-inhibitory effect of retinoic acid by promoting apoptosis in human breast cancer cells, *Mol Cell Biol* **16**:1138-49

Moras D and Gronemeyer H (1998) The nuclear receptor ligand-binding domain: structure and function *Curr Opin Cell Biol.* **10**:384-91

Paige LA, Christensen DJ, Gron H, Norris JD, Gottlin EB, Padilla KM, Chang CY, Ballas LM,

Hamilton PT, McDonnell DP, Fowlkes DM. (1999) Estrogen receptor (ER) modulators each induce distinct conformational changes in ER α and ER β. *Proc Natl Acad Sci U S A.* **96**:3999-4004

Rizvi NA, Marshall JL, Ness E, Yoe J, Gill GM, Truglia JA, Loewen GR, Jaunakais D, Ulm EH, Hawkins MJ. (1998) Phase I study of 9-cis-retinoic acid (ALRT1057 capsules) in adults with advanced cancer *Clin Cancer Res.* **4**:1437-42

Thacher SM, Vasudevan J, Chandraratna RA (2000) Therapeutic Applications for Ligands of Retinoid Receptors *Curr Pharm Des* **6**:25-58

Widschwendter M, Berger J, Daxenbichler G, Muller-Holzner E, Widschwendter A, Mayr A, Marth C, Zeimet AG (1997) Loss of retinoic acid receptor β expression in breast cancer and morphologically normal adjacent tissue but not in the normal breast tissue distant from the cancer *Cancer Res* **57**:4158-61

Appendices

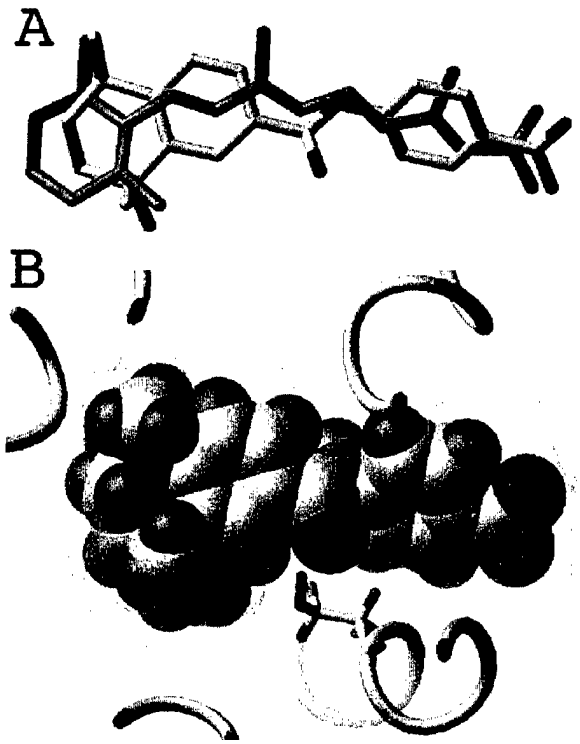


Figure 1: Docking of a known RARα specific agonist. The RARα selective agonist Am580 was docked into the modeled ligand binding pocket of RARα. A: The ligand (white sticks) superimposes with the crystal structure of bound all-trans RA (green). Hydrogens are not shown for clarity. B: Am580 (CPK display) fits tightly into the receptor's pocket (yellow wire), except for a ketone oxygen, which does not clash with the receptor only because it can share an hydrogen with Ser234 (displayed as stick). In RARβ and γ, Ser234 is replaced by an alanine which clashes with the ketone oxygen. The receptor in the vicinity of the ligand is shown as a white ribbon. Carbons, hydrogens, oxygens and nitrogen are colored white, grey, red and blue respectively.

Figure 2: Models of RAR LBD where the C-terminal helix H12 (green, magenta) is relocated onto different docking sites of the receptor (surface representation) could provide novel binding pocket conformations for the structure based discovery of agonists (cyan) with improved pharmacological profile.

Ligand	Activity	Score1	Score2	Score3	Score4	Selected	Binding
AGN193836	RARα agonist	-37.2	-36.5	-36.7	-35.3	+	+
ATRA	RAR pan-agonist	-51.7	-52.6	-51.8	-52.0	+	+
Ro415253	RARα antagonist	-28.9	-24.4	-39.0	-46.6	+	+
MX781	RAR antagonist	-45.3	-48.0	-40.2	-45.6	+	+
CD2366	RAR pan-antagonist.	-50.7	-50.8	-29.3	-29.3	+	+
Targretin	RXR pan-agonist	-25.4	-23.0	-22.2	-31.0	-	-
SR11203	RXR pan-agonist	-28.2	-22.7	-22.1	-27.5	-	-
Tamoxifen	ER modulator	-26.4	-24.6	-30.3	-23.4	-	-
Raloxifene	ER modulator	-15.6	-23.7	-18.4	-17.4	-	-
RU486	Progest Rec antag.	-21.4	-20.6	-20.3	-20.1	-	-
9cisRA	RAR/RXR agonist.	-38.8	-39.5	-33.5	-38.7	+	+
AGN193109	RAR pan-antagonist.	-55.1	-55.5	-41.2	-54.8	+	+
AGNpartia	RAR partial agonist.	-61.4	-61.3	-61.4	-61.0	+	+
Am580	RARα agonist	-46.6	-47.2	-46.6	-46.5	+	+
EM652	ER antagonist	-26.3	-23.1	-23.7	-27.3	-	-
Antagonist1	Novel RAR antag	-32.1	-32.1	-31.7	-31.6	+	+
Antagonist2	Novel RAR antag	-33.3	-29.7	-33.8	-33.8	+	+

Table 1: Control screening of known ligands. A similar screening as the one performed on the large library was carried out four times on a small database made of known nuclear receptor ligands: the compounds were automatically docked in the structure of RAR. The molecules which met at least once with the criteria for selection are listed as "Selected". The ligands which are experimentally binding to RAR are listed as "Binding".

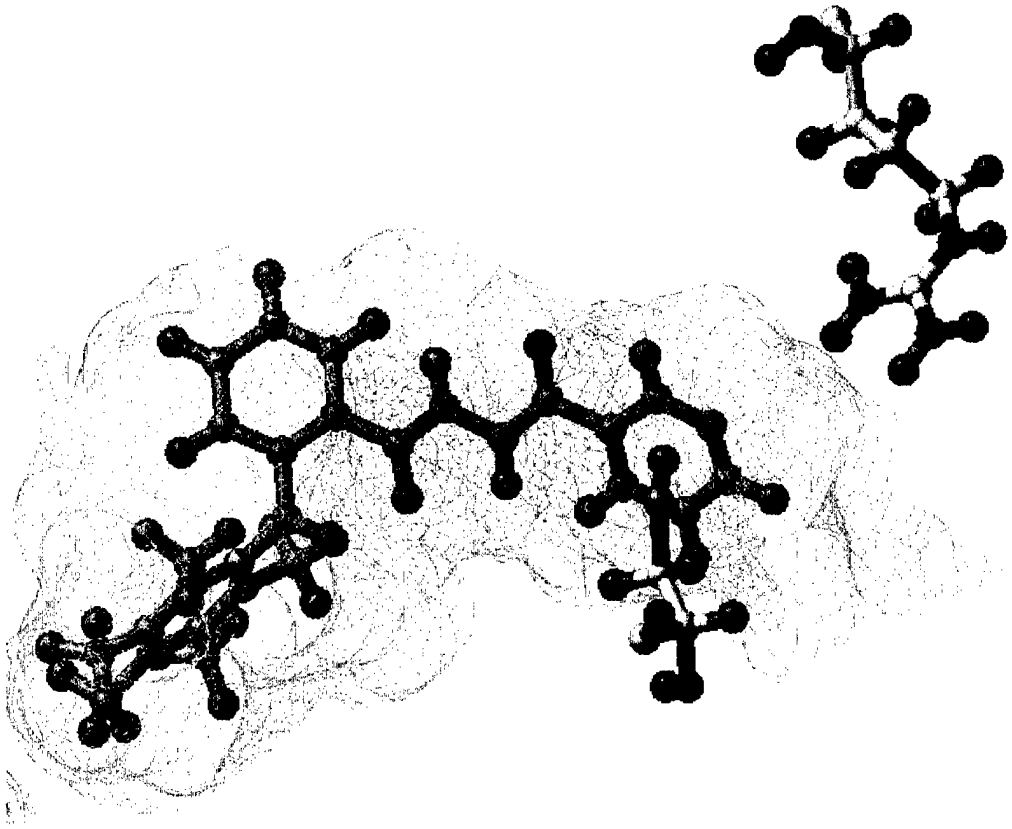


Figure 3: Unlike all retinoid receptor ligands described in the literature, Agonist 3 (green) does not have a carboxylate, but a pyridine which makes a thiolate salt with a cystein (white) of the receptor, and could also interact with a neighboring arginine (white). This compound is active at 200 nM, and the rational design of analogs, which conserve this original feature could result in leads against breast cancer with improved specificity and toxicity profiles. We embarked on the development of such molecules in collaboration with renowned chemists and biologists.

New York University
A private University in the Public Service

Molecular Synthesis and Design Group
Stephen R. Wilson
Department of Chemistry
New York University
New York, NY 10003
Ph: 212-998-8461
Fax: 212-260-7905
steve.wilson@nyu.edu

May 24, 2000

Matthieu Schapira, Ph.D.
Skirball Institute, lab 3/7
NYU Medical Center
540 First Ave., New York, NY10016

Dear Matthieu:

This letter is to confirm my interest and enthusiasm for continuing our collaborations involving synthesis and combinatorial chemistry support of your projects. As you know, my laboratory in the Chemistry Department at NYU is active in molecule synthesis and design. Most recently we have been involved in developing chemistry of RAR agonist 3 with your group.

I am looking forward to continuing to work with you.

Regards,

Stephen R. Wilson
Professor of Chemistry

NYU School of Medicine

NEW YORK UNIVERSITY SCHOOL OF MEDICINE

A private university in the public service

550 FIRST AVENUE, NEW YORK, NY 10016
CABLE ADDRESS: NYUMEDIC

Department of Medicine
Division of Clinical and Molecular Endocrinology

(212) 263-6279 (Voice)
(212) 263-7701 (FAX)

Herbert H. Samuels, M.D.
Director

Manfred Blum, M.D.
Veronica Catanese, M.D.
Raquel Cintron
Vandana Desai-Yajnik, Ph.D.
Stuart Weiss, M.D.
William Drucker, M.D.
Loren Greene, M.D.
Sandra Kammerman, M.D.
David Kleinberg, M.D.
Brian Levy, M.D.
Mary McCarthy
Connie Newman, M.D.
Valerie Peck, M.D.
Bruce Raaka, Ph.D.
Stephen Richardson, M.D.
Ysa Romero
David Ron, M.D.
Arnold Saperstein, M.D.
Terry Seltzer, M.D.
Fredrick Stanley, Ph.D.

May 23, 2000

Matthieu Schapira, Ph.D.
Skirball Institute
NYU School of Medicine
540 First Ave., New York, NY 10016

Dear Matthieu:

This is a letter in support of our continuing collaborative studies on the *in silico* identification of novel retinoid receptor agonists and antagonists. We are delighted to continue our collaboration and will, in addition to other interesting ligand candidates, analyze the biological activity of synthetic analogs of RAR agonist 3.

All the best,

Herbert H. Samuels, M.D.
Professor of Medicine and Pharmacology
Head, Division of Clinical and
Molecular Endocrinology

MOLECULAR AND CELLULAR BIOLOGY, Oct. 1999, p. 7191–7202
0270-7306/99/$04.00+0

Vol. 19, No. 10

NRIF3 Is a Novel Coactivator Mediating Functional Specificity of Nuclear Hormone Receptors

DANGSHENG LI,[1] VANDANA DESAI-YAJNIK,[1] ERIC LO,[1] MATTHIEU SCHAPIRA,[2] RUBEN ABAGYAN,[2] AND HERBERT H. SAMUELS[1]*

Division of Molecular Endocrinology, Departments of Medicine and Pharmacology,[1] and Structural Biology, Skirball Institute of Biomolecular Medicine,[2] New York University School of Medicine, New York, New York 10016

Received 25 February 1999/Returned for modification 5 April 1999/Accepted 16 July 1999

Many nuclear receptors are capable of recognizing similar DNA elements. The molecular event(s) underlying the functional specificities of these receptors (in regulating the expression of their native target genes) is a very important issue that remains poorly understood. Here we report the cloning and analysis of a novel nuclear receptor coactivator (designated NRIF3) that exhibits a distinct receptor specificity. Fluorescence microscopy shows that NRIF3 localizes to the cell nucleus. The yeast two-hybrid and/or in vitro binding assays indicated that NRIF3 specifically interacts with the thyroid hormone receptor (TR) and retinoid X receptor (RXR) in a ligand-dependent fashion but does not bind to the retinoic acid receptor, vitamin D receptor, progesterone receptor, glucocorticoid receptor, or estrogen receptor. Functional experiments showed that NRIF3 significantly potentiates TR- and RXR-mediated transactivation in vivo but has little effect on other examined nuclear receptors. Domain and mutagenesis analyses indicated that a novel C-terminal domain in NRIF3 plays an essential role in its specific interaction with liganded TR and RXR while the N-terminal LXXLL motif plays a minor role in allowing optimum interaction. Computer modeling and subsequent experimental analysis suggested that the C-terminal domain of NRIF3 directly mediates interaction with liganded receptors through an LXXIL (a variant of the canonical LXXLL) module while the other part of the NRIF3 protein may still play a role in conferring its receptor specificity. Identification of a coactivator with such a unique receptor specificity may provide new insight into the molecular mechanism(s) of receptor-mediated transcriptional activation as well as the functional specificities of nuclear receptors.

Nuclear hormone receptors are ligand-regulated transcription factors that play diverse roles in cell growth, differentiation, development, and homeostasis. The nuclear receptor superfamily has been divided into two subfamilies: the steroid receptor family and the thyroid hormone/retinoid (nonsteroid) receptor family (51). The steroid receptor family includes receptors for glucocorticoids (GR), mineralcorticoids, progestins (PRs), androgens (AR), and estrogens (ERs) (51). The nonsteroid receptor family includes receptors for thyroid hormones (TRs), retinoids (retinoic acid receptors [RARs] and retinoid X receptors [RXRs]), 1,25-$(OH)_2$-vitamin D (VDR), and prostanoids (peroxisome proliferator-activated receptors [PPARs]) as well as many orphan receptors whose ligands (if any) remain to be defined (49, 51). Members of the nuclear receptor superfamily have common structural and functional motifs. Nevertheless, an important difference exists between the two subfamilies. Steroid receptors act primarily as homodimers by binding to their cognate palindromic hormone response elements (HREs) (77, 78). In contrast, members of the nonsteroid receptor family can bind to DNA as monomers, homodimers, and heterodimers (25, 78). Their corresponding HREs are also complex and can be organized as direct repeats, inverted repeats, and everted repeats (49). Therefore, the combination of heterodimerization and HRE complexity provides the potential of generating enormous diversity in receptor-mediated regulation of target gene expression.

Structural and functional studies indicate that the ligand binding domain (LBDs) of many members of the thyroid hormone/retinoid receptor family harbor diverse functions. In addition to binding to ligands, the LBD also plays roles in mediating receptor dimerization, hormone-dependent transactivation, and, with TR and RAR, ligand-relieved gene silencing (54, 61). The carboxyl-terminal helix of the LBD has been implicated in playing an important role in ligand-dependent conformational changes and transactivation (6, 9, 21, 43). Although it has been suggested that an activation function (AF-2) resides in this C-terminal helix, recent studies indicate that AF-2 results from a ligand-induced conformational change involving diverse areas of the LBD (23, 66). Thus, ligand binding serves to switch the receptor from one functional state (e.g., inactive or silencing) to another (e.g., transactivation).

Although much has been learned from studying the structures and functions of these receptors, the detailed molecular mechanism(s) of transcriptional regulation by these receptors is not well understood. Efforts to understand the molecular mechanism of transcriptional repression by unliganded TRs and RARs have led to the description (12) and isolation of the putative corepressor proteins SMRT and N-CoR, which interact with the LBDs of these receptors in the absence of their ligands (15, 36). The recent discovery that both SMRT and N-CoR form complexes with Sin3 and a histone deacetylase suggests that chromatin remodeling by histone deacetylation may play a role in receptor-mediated transcriptional repression (33, 55).

In a somewhat parallel approach, the identification of coactivators has recently received extensive experimental attention in order to elucidate the molecular mechanism(s) of transcriptional activation by nuclear receptors (27). Identified coactiva-

* Corresponding author. Mailing address: Division of Molecular Endocrinology, Departments of Medicine and Pharmacology, New York University School of Medicine, 550 First Ave., New York, NY 10016. Phone: (212) 263-6279. Fax: (212) 263-7701. E-mail: samueh01@mcrcr.med.nyu.edu.

tor proteins primarily belong to two groups: the SRC-1 family and the CREB-binding protein (CBP)/p300 family. The SRC-1 family includes SRC-1/NCoA-1 (37, 58, 74) and the related proteins GRIP1/TIF2/NCoA-2 (34, 35, 74, 79) and AIB1/p/CIP/ACTR/RAC3/TRAM-1 (2, 14, 44, 73, 74). The second group of coactivators includes CBP and its homolog p300, which not only influence the activities of nuclear receptors (13, 31, 37) but also functionally interact with many transcription factors such as CREB (3, 16, 40, 46), the Stats (10, 87), AP1 (4, 7), and p53 (28, 45). There are also coactivator proteins that do not belong to these two groups, such as ARA70 (85), PGC-1 (60), and the recently reported RNA coactivator SRA (41). Members of both the SRC-1 family and CBP/p300 family have been shown to possess histone acetyltransferase activities (8, 14, 57, 69), suggesting that chromatin remodeling by histone acetylation is an important mechanism involved in transcriptional activation by ligand-bound nuclear receptors.

Interaction of members of the SRC-1 and CBP/p300 families with nuclear receptors occurs through conserved LXXLL motifs (32), which interact with a hydrophobic cleft in the receptor LBD formed as a result of conformational changes mediated by ligand binding (19, 23, 56). SRC-1/NCoA-1 and GRIP1/TIF2 contain three LXXLL regions or boxes (referred to as LXDs or nuclear receptor boxes) that differentially interact with nuclear receptors so that different nuclear receptors functionally utilize different LXXLL boxes (19, 52). Thus, ER uses the second LXXLL box of SRC-1/NCoA-1 while PR uses both the first and second LXXLL boxes for optimal interaction. In contrast, TR and RAR require both the second and third LXXLL boxes for optimal interaction (52). The specificities of receptor recognition by the different LXXLL boxes of SRC-1/NCoA-1 are primarily mediated by 8 amino acid residues C terminal to the LXXLL motif rather than by the 2 amino acids (XX) within the motif itself. Thus, while members of the SRC-1 family are capable of interacting with many nuclear receptors, the molecular details of such interactions differ for each receptor in the number or combination of LXXLL boxes used as well as in the critical amino acid residues surrounding the LXXLL motifs.

While much has been learned from the study of known coactivators, a number of key mechanistic questions remain to be answered. For example, many nuclear receptors can recognize common DNA elements, (25, 49, 51), while not all are capable of regulating genes containing those elements (20, 47, 65). Thus, how native target genes containing such elements are selectively regulated by specific receptors is a very important but poorly understood problem. Although the various LXXLL boxes of SRC-1 and GRIP1 show differential receptor preference (19, 52), these coactivators are unlikely to play a primary role in mediating effects that are receptor specific since they appear to interact with all ligand-bound nuclear hormone receptors. Thus, the detailed molecular mechanism(s) underlying receptor-specific regulation of gene expression remains to be elucidated. Whether a coactivator(s) contributes to this specificity is currently unknown.

To further our understanding of the molecular events underlying receptor-activated transcription, we sought to identify additional coactivators using a yeast two-hybrid screening strategy (29). In this paper, we report the isolation of a novel coactivator for nuclear receptors, designated NRIF3 (for nuclear receptor-interacting factor 3). Fluorescence microscopy indicates that NRIF3 is a nuclear protein. The yeast two-hybrid and in vitro binding assays revealed that NRIF3 interacts specifically with TR and RXR in a ligand-dependent fashion but does not interact with other examined nuclear receptors. Transfection experiments indicated that NRIF3 selectively potentiates TR- and RXR-mediated transactivation in vivo. The NRIF3 gene encodes a small protein of 177 amino acids and, other than having an N-terminal LXXLL motif, has no homology with known coactivator genes. The results of a combination of computer modeling and domain and mutagenesis analyses suggest that NRIF3 interacts with nuclear receptors through its C-terminal domain that contains a novel LXXIL module while another part of NRIF3 may contribute to its observed receptor specificity. These findings may provide novel insight into the molecular mechanism(s) of receptor-mediated transcriptional activation as well as the functional specificities of nuclear receptors.

MATERIALS AND METHODS

Isolation of NRIFs and the yeast two-hybrid assay. The Brent two-hybrid system (29) was employed to isolate candidate cDNA clones interacting with LexA-TRα in a ligand-dependent fashion. Full-length chicken TRα (cTRα) was fused in frame to the C terminus of the LexA DNA binding domain (DBD) in pEG202 (29). The LexA-TRα bait, the LacZ reporter (pSH18-34), and a pJG4-5-based HeLa cell cDNA library were transformed into the yeast strain EGY48 (29). The transformants were selected on Gal–Raf–X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) medium in the absence of leucine and were further screened for the expression of LacZ in the presence of 1 μM T3. Blue colonies were picked and reexamined for T3-dependent expression of LacZ. Positive yeast clones were then selected, and plasmids harboring candidate prey cDNAs were isolated. An individual candidate prey plasmid was then amplified in *Escherichia coli* and retransformed into the original yeast strain to confirm the interaction phenotype. The cDNA inserts were then sequenced with an automatic sequencer. Four novel clones (NRIF1, -2, -3, and -4) were obtained. Among them, NRIF3 was a full-length clone.

Wild-type NRIF3, the β3-endonexin long form (EnL) and short form (EnS), and the L9A NRIF3 mutant protein were examined for their interaction with various nuclear receptors in a yeast two-hybrid assay. The following receptor baits were used: the LexA-cTRα LBD, LexA-human TRβ (hTRβ) LBD, LexA-hRARα LBD, LexA-hRXRα LBD, and LexA-hGR LBD. The NRIF3 C-terminal domain (NCD) was fused in frame with the LexA DBD and examined for interaction with receptor LBDs with the following preys: the B42-cTRα LBD, B42-hRARα LBD, and B42-hRXRα LBD expressed from pJG4-5. Yeast cells harboring appropriate plasmids were grown in selective media with Gal-Raf in the presence or absence of cognate ligand (1 μM T3 for TR, all *trans* or 9-*cis* RA for RAR, 9-*cis* RA for RXR, and 10 μM deoxycorticosterone for GR) overnight before β-galactosidase activity was assayed with *o*-nitrophenyl β-D-galactopyranoside as the substrate. β-Galactosidase units were calculated with the formula $(OD_{420} \times 1{,}000)/(\text{minutes of incubation} \times OD_{600}$ of yeast suspension), where OD_{420} and OD_{600} are the optical densities at 420 and 600 nm, respectively.

Fluorescence microscopy. Full-length NRIF3 was cloned into the green fluorescent protein (GFP) fusion protein expression vector pEGFP (Clontech). The resulting GFP-NRIF3 vector and the control plasmid pEGFP were transfected into HeLa cells by calcium phosphate coprecipitation. Cells were incubated at 37°C for 24 h before the examination with a fluorescence microscope to determine the subcellular location of GFP-NRIF3 or the GFP control.

In vitro binding assay. Full-length NRIF3 was cloned into pGEX2T, a bacterial glutathione *S*-transferase (GST) fusion protein expression vector (Pharmacia). The GST-NRIF3 fusion protein was expressed in *E. coli* and affinity purified with glutathione-agarose beads (30). ^{35}S-labeled full-length cTRα, hRARα, hRXRα, hVDR, hGR, hPR, and hER were generated by in vitro transcription and translation with a reticulocyte lysate system (Promega). Binding was performed as previously described (30) with the following buffer: 20 mM HEPES (pH 7.9)–1 mM $MgCl_2$–1 mM dithiothreitol–10% glycerol–0.05% Triton X-100–1 μM $ZnCl_2$–150 mM KCl. Appropriate ligands were added into the binding reaction mixture where indicated in the figures in the following concentrations: 1 μM T3 for TR, 1 μM all-*trans* RA or 9-*cis* RA for RAR, 1 μM 9-*cis* RA for RXR, and 150 nM 1,25-$(OH)_2$-vitamin D_3, dexamethasone, progesterone, or estradiol for VDR, GR, PR, or ER, respectively. After the binding reaction, the beads were washed three times and the labeled receptors bound to the beads were examined by sodium dodecyl sulfate–10% polyacrylamide gel electrophoresis followed by autoradiography. Five percent of the ^{35}S-labeled receptor input was also electrophoresed in the same gel.

Transfection studies. Most reporters used in this study, including IR-ΔMTV-CAT, DR4-ΔMTV-CAT, GH-TRE-tk-CAT, and IR+3 (ERE)-ΔMTV-CAT, have been described previously (5, 25, 78). A DR1-ΔMTV-CAT reporter responsive to RXR was obtained from Ron Evans. A GRE/PRE-tk-CAT reporter was obtained from Gunther Schutz. (IR)2-TATA-CAT was constructed in our laboratory by cloning two copies of the inverted-repeat (IR) sequence (AGG TCA TGACCT) upstream of a TATA element derived from the thymidine kinase (tk) promoter. An hVDR expression vector and VDRE-ΔMTV-CAT containing the VDRE from the osteocalcin promoter were obtained from J. Wesley Pike. Vectors expressing cTRα, hRARα, hRXRα, rat GR (rGR), hPR,

and hER have been described previously (17, 25, 26, 50, 53, 81). The NRIF3 expression vector was constructed by cloning full-length NRIF3 into a pExpress vector (25). Appropriate plasmids were transfected into HeLa cells by calcium phosphate coprecipitation with 25 to 100 ng of the receptors, 250 to 500 ng of the chloramphenicol acetytransferase (CAT) reporters, and 750 ng of the NRIF3 or control pExpress vector. After transfection, cells were incubated at 37°C (with or without cognate ligands) for 42 h before being harvested. CAT assays were carried out as previously described (30). Relative CAT activity was determined as the percent acetylation of the substrate per 30 μg of cell protein in a 15-h incubation at 37°C. The results were calculated from duplicate or quadruplicate samples, and the variation among samples was less than 10%.

Domain and mutagenesis analyses. To construct pJG4-5-derived vectors expressing EnL or EnS, the pJG4-5/NRIF3 plasmid was digested with *Nco*I and *Xho*I and the resulting vector fragment was gel purified. This fragment was then ligated to an EnL or EnS insert generated from pExpress-EnL or pExpress-EnS by *Nco*I/*Sal*I double digestion. The resulting pJG4-5/EnL or pJG4-5/EnS plasmid was confirmed by sequence analysis. The L9A mutant form of NRIF3 was generated by site-directed mutagenesis by a PCR-based method, and the mutation was confirmed by sequence analysis. pJG4-5-derived vectors expressing EnL, EnS, or the L9A NRIF3 mutant form were transformed into yeast strains harboring the LacZ reporter (pSH18-34) and appropriate bait plasmids (LexA-TR, LexA-RAR, LexA-RXR, and LexA-GR). Transformants were subjected to quantitative assays of β-galactosidase activity as described above.

To construct the bait plasmid expressing LexA-NCD, a derivative of pEG202 (which contains a new polylinker) was digested with *Nco*I and *Xho*I and ligated to synthetic oligonucleotides that encode the last 16 amino acids of NRIF3 (residues 162 to 177). Similarly, the mutant NCD was generated by using oligonucleotides that contain the designed mutations in the ligation reaction. Both constructs were confirmed by sequence analysis. Bait plasmids expressing LexA-NCD or LexA-mutant NCD were transformed together with one of the following prey plasmids, B42-TR LBD, B42-RXR LBD, or B42-RAR LBD, into the yeast strain that harbors the LacZ reporter (pSH18-34). Subsequent two-hybrid assays were carried out as described above.

Docking of coactivator peptides to receptors. We built a model of the interaction between the 17-residue C-terminal peptide of NRIF3 (KASRHLDSYE FLKAILN) and the LBDs of several receptors (TRα was used as an example in the experiment reflected in Fig. 10). An LXXIL motif within the NRIF3 peptide is underlined. A similar modeling procedure was carried out on a 20-residue peptide (SLTERHKILHRLLQEGSPSD) of the second LXXLL box of SRC-1 (52). We hypothesized that the LXXIL motif of the C terminus of NRIF3 contacts the coactivator binding site of the nuclear receptors, and the automatic docking procedure was carried out towards this site (71, 75, 76). Two critical features of the interaction between the LBDs of nuclear hormone receptors and their coactivators were used to build the models. (i) One was the "charge clamp," initially observed in the complex between SRC-1 and PPARγ (56), where a conserved glutamate and lysine at opposite ends of the hydrophobic cavity of the receptors contact the backbone of the coactivator's LXXLL box. This feature enabled us to orient the NRIF3 helical peptide. (ii) The other feature was that the leucines of the LXXLL motif of SRC-1 are buried in the hydrophobic cavity of the receptor. This feature allowed us to predict the side of the NRIF3 peptide which faces the receptor.

The coactivator peptides were assigned a helical secondary structure, the backbone φ and ψ angles being −62 and −41 degrees, respectively. The ω angle was set to 180 degrees. Loose distance restraints were set between the charge clamp of the receptors (56) and C^{α} atoms of the peptide. The energy of the complex was minimized in the internal coordinate space by using the modified ECEPP/3 potentials. The subset of the variables minimized by the ICM method (1, 71, 76) included the side chains of the receptor, six positional variables of the helix, and the side chain torsion angles of the helix.

Binding energy calculation. The binding energy was calculated by the partitioning method as described elsewhere (64). Briefly, the binding energy function is partitioned into three terms: the surface (or hydrophobic) term, determined as the product of the solvent-accessible surface by a surface tension of 30 cal/mol/$Å^2$; the electrostatic term, calculated by a boundary element algorithm, with a dielectric constant of 8; and the entropic term, which results from the decrease in conformational freedom of residue side chains partially or completely buried upon complexation.

RESULTS

Cloning of NRIF3 cDNA. To isolate potential coactivators mediating the transcriptional activation functions of nuclear receptors, we employed a yeast two-hybrid screening strategy (29). A bait expressing a full-length TRα fused to the C terminus of the LexA DBD was used to screen a HeLa cell cDNA library cloned into pJG4-5 (29). Candidate clones that exhibited a thyroid hormone (T3)-dependent interaction with LexA-TRα were selected and further examined and sequenced. Four novel clones were identified, and all were found to exhibit

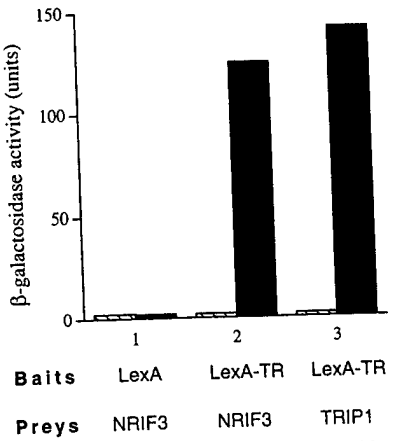


FIG. 1. Hormone-dependent interaction of NRIF3 with the LBD of TR. Induction of β-galactosidase activity by thyroid hormone (T3) was measured in the yeast strain EGY48 transformed with a bait vector expressing the LexA-cTRα LBD and a prey plasmid expressing NRIF3 fused to the B42 activation domain (29). The bait LexA alone was used as the negative control. The prey B42–Trip1 was used as the positive control. Hatched bars, without T3; filled bars, with 1 μM T3.

levels of interaction with the LBD of TRα similar to the levels they exhibited with the full-length receptor (data not shown). These clones were designated NRIF1, -2, -3, and -4. Not surprisingly, the LBD of TRβ was also found to interact with these NRIFs in a T3-dependent manner (data not shown). Among these four isolated NRIFs, NRIF3 was a full-length clone. As shown in Fig. 1, LexA alone (negative control) did not interact with NRIF3 (as indicated by the low β-galactosidase activity) and incubation with T3 had no effect. Similarly, no interaction was detected between the LexA-TR LBD and B42 alone with or without T3 (data not shown). The LexA-TR LBD also showed little interaction with NRIF3 in the absence of T3. However, incubation with T3 resulted in strong stimulation of the NRIF3-TR LBD interaction (Fig. 1). The extent of T3-dependent interaction between NRIF3 and the LexA-TR LBD was similar to that of Trip1 (Fig. 1), one of the first TR-interacting factors cloned in a two-hybrid screen (42).

Sequence analysis of NRIF3. Sequence analysis of the NRIF3 cDNA revealed a single open reading frame encoding a polypeptide of 177 amino acids (Fig. 2). NRIF3 has no homology with members of the SRC-1 and CBP/p300 families. The size of NRIF3 is in sharp contrast to the size of CBP/p300 (around 300 kDa) or of SRC-1 family members (around 160 kDa). NRIF3 contains a putative nuclear localization signal (KRKK), as well as one copy of an LXXLL motif (amino acids 9 to 13) that was recently identified as being essential for the interaction of a number of putative coactivators with nuclear receptors (32).

A database search identified two highly related homologs of NRIF3, which were previously designated β3-endonexin short and long forms (67). The endonexin short form (EnS) was originally isolated from a two-hybrid screen intended to clone factors that interact with the cytoplasmic tail of integrin β3 (67). The long form (EnL) was then identified as an alternatively spliced product of the same gene. However, the long form does not bind to integrin β3 (67). Nucleotide sequence comparisons between cDNAs of NRIF3 and EnS or EnL indicate that NRIF3 is a third alternatively spliced product of the same gene (alignment not shown). The precise function(s) of the two endonexin proteins is under investigation (reference 66a and see Discussion).

```
1   CAGCGGCAGTGGTGCTTTCCCGAATCTCAGAATGCCTGTTAAAAGATCACTGAAGTTGGA
                                M  P  V  K  R  S  L  K  L  D   10

61  TGGTCTGTTAGAAGAAAATTCATTTGATCCTTCAAAAATCACAAGGAAGAAAAGTGTTAT
      G  L  L  E  E  N  S  F  D  P  S  K  I  T  R  K  K  S  V  I   30

121 AACTTATTCTCCAACAACTGGAACTTGTCAAATGAGTCTATTTGCTTCTCCCACAAGTTC
      T  Y  S  P  T  T  G  T  C  Q  M  S  L  F  A  S  P  T  S  S   50

181 TGAAGAGCAAAAGCACAGAAATGGACTATCAAATGAAAAGAGAAAAAAATTGAATCACCC
      E  E  Q  K  H  R  N  G  L  S  N  E  K  R  K  K  L  N  H  P   70

241 CAGTTTAACTGAAAGCAAAGAATCTACAACAAAAGACAATGATGAATTCATGATGTTGCT
      S  L  T  E  S  K  E  S  T  T  K  D  N  D  E  F  M  M  L  L   90

301 ATCAAAAGTTGAGAAATTGTCAGAAGAAATCATGGAGATAATGCAAAATTTAAGTAGTAT
      S  K  V  E  K  L  S  E  E  I  M  E  I  M  Q  N  L  S  S  I  110

361 ACAGGCTTTGGAGGGCAGTAGAGAGCTTGAAAATCTCATTGGAATCTCCTGTGCATCACA
      Q  A  L  E  G  S  R  E  L  E  N  L  I  G  I  S  C  A  S  H  130

421 TTTCTTAAAAAGAGAAATGCAGAAAACCAAAGAACTAATGACAAAAGTGAATAAACAAAA
      F  L  K  R  E  M  Q  K  T  K  E  L  M  T  K  V  N  K  Q  K  150

481 ACTGTTTGAAAAGAGTACAGGACTTCCTCACAAAGCATCACGTCATCTTGACAGCTATGA
      L  F  E  K  S  T  G  L  P  H  K  A  S  R  H  L  D  S  Y  E  170

541 ATTCCTTAAAGCCATTTTAAACTGAGGCATTAAGAAGAAATGCACTCACCATGAGCACCA
      F  L  K  A  I  L  N  *
```

FIG. 2. Nucleotide and deduced amino acid sequences of NRIF3. Only part of the cDNA sequence is shown. A putative nuclear localization signal (KRKK) is underlined. The putative LXXLL motif is shown with a double underline. NRIF3 and EnL have 95% identity. They differ only in their C termini, where the last 16 amino acids (dotted underline) in NRIF3 are replaced with 9 different amino acids (GQPQMSQPL) in EnL. EnS consists of 111 amino acids and is 100% identical to the first 111 amino acids of NRIF3 or EnL.

NRIF3 localizes to the cell nucleus. Although a putative nuclear localization signal was found in NRIF3, we considered it important to identify the subcellular location of the NRIF3 protein since extensive homology was found between NRIF3 and the two endonexins. The entire NRIF3 open reading frame was fused to the C terminus of GFP (18). The resulting GFP-NRIF3 fusion protein was expressed in HeLa cells by transient transfection, and the subcellular location of the fusion protein was visualized by fluorescence microscopy. As shown in Fig. 3, the control GFP protein distributed throughout the cell while GFP-NRIF3 localized exclusively to the nucleus. This result suggests that NRIF3 is a nuclear protein, which is compatible with its putative role as a nuclear receptor coactivator.

Selective interaction of NRIF3 with liganded nuclear receptors in yeast. Although NRIF3 was originally cloned with full-length TRα as the bait, we later identified that the region of the receptor responsible for NRIF3 binding is its LBD (Fig. 1). A common feature among most of the known coactivators that show ligand-dependent interaction with nuclear receptors is the presence of the LXXLL motif(s) in their receptor interaction domains. The LXXLL motif appears to be involved in direct contact with a structurally conserved surface in the ligand-bound LBDs of the receptors (23), which may provide

FIG. 3. NRIF3 is a nuclear protein. HeLa cells were transfected with an expression vector for GFP (left panel) or GFP-NRIF3 (right panel). The cellular location of the expressed proteins was visualized by fluorescence microscopy.

TABLE 1. Interaction of NRIF3 with nuclear receptors in yeast[a]

Bait	Prey	β-Galactosidase activity		
		Without ligand	With ligand	Fold stimulation
LexA	NRIF3-B42	2.3	1.9	0.8
LexA-TR	NRIF3-B42	1.8	125	69
LexA-RAR	NRIF3-B42	0.1	0.1	1
LexA-RXR	NRIF3-B42	0.2	63	315
LexA-GR	NRIF3-B42	0.8	0.6	0.8

[a] The LacZ reporter activities were determined for yeast strains harboring the indicated bait and prey plasmids in the presence or absence of cognate ligands as described in Materials and Methods. See the text for detailed explanations.

the molecular basis for the broad spectrum of receptor binding by coactivators such as SRC-1 and GRIP1. Since a putative LXXLL motif is also present in NRIF3 (amino acids 9 to 13), we asked whether NRIF3 also interacts with the LBDs of other nuclear receptors.

The LBDs of several nuclear receptors were examined for interaction with NRIF3 in a yeast two-hybrid assay. As shown in Table 1, NRIF3 does not interact with LexA alone (negative control) with or without ligand. LexA-TR and LexA-RXR showed little (if any) interaction with NRIF3 in the absence of their cognate ligands. However, the presence of T3 (for TR) or 9-*cis* RA (for RXR) resulted in a strong stimulation of their interaction with NRIF3, as indicated by the induction of β-galactosidase activity (Table 1). Interestingly, when LexA-RAR or LexA-GR was used as the bait, no interaction was detected with NRIF3 in the presence or absence of their cognate ligands (Table 1). The finding that NRIF3 interacts with TR but not RAR was surprising in light of a recent study which showed that TR and RAR functionally interact with the same LXXLL boxes (boxes 2 and 3) of SRC-1/NCoA-1 (52). As positive controls, we confirmed that both LexA-RAR and LexA-GR exhibited ligand-dependent interaction with other coactivators that are not receptor specific (data not shown). Taken together, these results suggest that NRIF3 exhibits differential specificities in its interactions with different nuclear receptors.

NRIF3 specifically binds to TR and RXR but not to other nuclear receptors in vitro. To further examine the interaction between NRIF3 and various nuclear receptors as well as to confirm the potential receptor specificity of NRIF3, in vitro GST binding assays were performed (30). ^{35}S-labeled nuclear receptor, generated by in vitro transcription and translation, was incubated with purified GST-NRIF3 or the GST control bound to glutathione-agarose beads. All binding assays were carried out with or without the cognate ligand of the examined receptor. As shown in Fig. 4 (top left), TR and NRIF3 interact poorly in the absence of T3. Addition of T3 resulted in a strong increase in TR binding to GST-NRIF3, confirming that NRIF3 associates with TR in a T3-dependent manner. Using similar binding assays, we also studied the interaction of NRIF3 with six other nuclear receptors. Consistent with our findings from the yeast two-hybrid experiments (Table 1), NRIF3 interacted with RXR in vitro in a ligand-dependent manner (Fig. 4) but showed little or no binding to other nuclear receptors (RAR, VDR, GR, PR, and ER) in the presence or absence of their cognate ligands (Fig. 4). Taken together, the results of the yeast two-hybrid (Table 1) and the in vitro binding (Fig. 4) assays suggest that NRIF3 possesses a distinct receptor specificity.

NRIF3 selectively potentiates TR- and RXR-mediated transactivation in vivo. To examine the potential role of NRIF3 in TR-mediated transactivation, transfection studies were carried

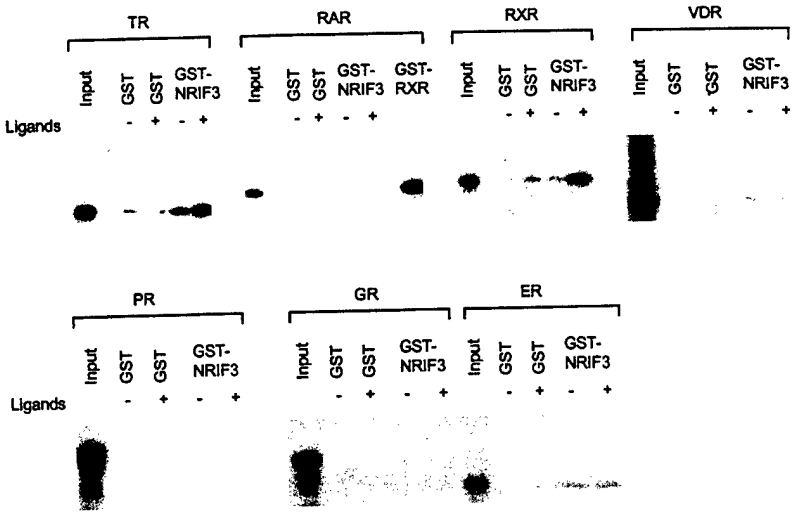


FIG. 4. Characterization of the NRIF3 interaction with nuclear receptors in vitro. A ^{35}S-labeled full-length receptor (cTRα, hRARα, hRXRα, hVDR, hPR, hGR, or hER) was incubated with an affinity-purified GST control or GST-NRIF3 linked to glutathione-agarose beads. The binding was performed in the absence (−) or presence (+) of cognate ligands as described in Materials and Methods. After incubation and washing, the bound receptors were analyzed by sodium dodecyl sulfate–10% polyacrylamide gel electrophoresis and detected by autoradiography. The input lane in each binding assay represents 5% of the total ^{35}S-labeled receptor used in each incubation. GST-RXR was used as a positive control for RAR binding.

out. HeLa cells, which lack endogenous TR (25), were transfected with a vector expressing TR and a CAT reporter under the control of the ΔMTV basal promoter linked to an idealized IR (AGGTCATGACCT) TRE sequence (IR-ΔMTV-CAT) (25), along with either a control plasmid or a vector expressing NRIF3. As shown in Fig. 5A, NRIF3 significantly enhances TR-mediated activation of the CAT reporter (typically 2.5- to 3-fold). As a control, we also examined the effect of CBP, a reported coactivator for nuclear receptors (13, 37), and found that its expression results in a degree of enhancement similar to that with NRIF3 (around threefold) (Fig. 5A).

We also examined another CAT reporter controlled by the herpesvirus tk promoter linked to native rat growth hormone TRE sequences (5). NRIF3 was found to also enhance TR-mediated activation of this reporter (about 3.5-fold) (Fig. 5B). In addition, using similar transfection assays, we found that NRIF3 enhances TR-mediated activation of two other reporters, (IR)2-TATA-CAT and DR4-ΔMTV-CAT (data not shown). Therefore, NRIF3 potentiates TR-mediated transactivation in a variety of different TRE and promoter contexts. Taken together, the results of these transfection studies suggest that NRIF3 can function as a coactivator of TR.

To examine whether NRIF3 can also act as a coactivator for RXR, HeLa cells were transfected with the IR-ΔMTV-CAT reporter, whose IR sequence can also function as a strong response element for the RXR(s) and RAR(s) (25, 49, 61). HeLa cells express the endogenous RXR(s) and RAR(s), as the activity of the IR-ΔMTV-CAT reporter was strongly stimulated by their cognate ligands, even without cotransfection of any receptor expression plasmid (Fig. 6A, bars 1, 3, and 5).

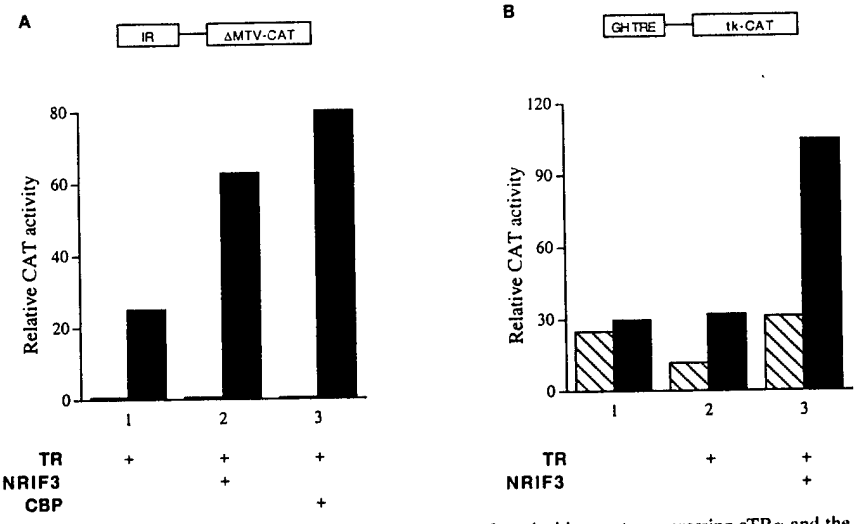


FIG. 5. NRIF3 enhances TR-mediated transactivation in vivo. HeLa cells were transfected with a vector expressing cTRα and the IR-ΔMTV-CAT reporter (A) or the GH-TRE-tk-CAT reporter (B) in the presence (filled bars) or the absence (hatched bars) of 1 μM T3. The vector expressing NRIF3 or the empty control vector was cotransfected to examine the effect of NRIF3 on TR-mediated activation. In panel A, the effect of CBP was compared to that of NRIF3. GH, growth hormone.

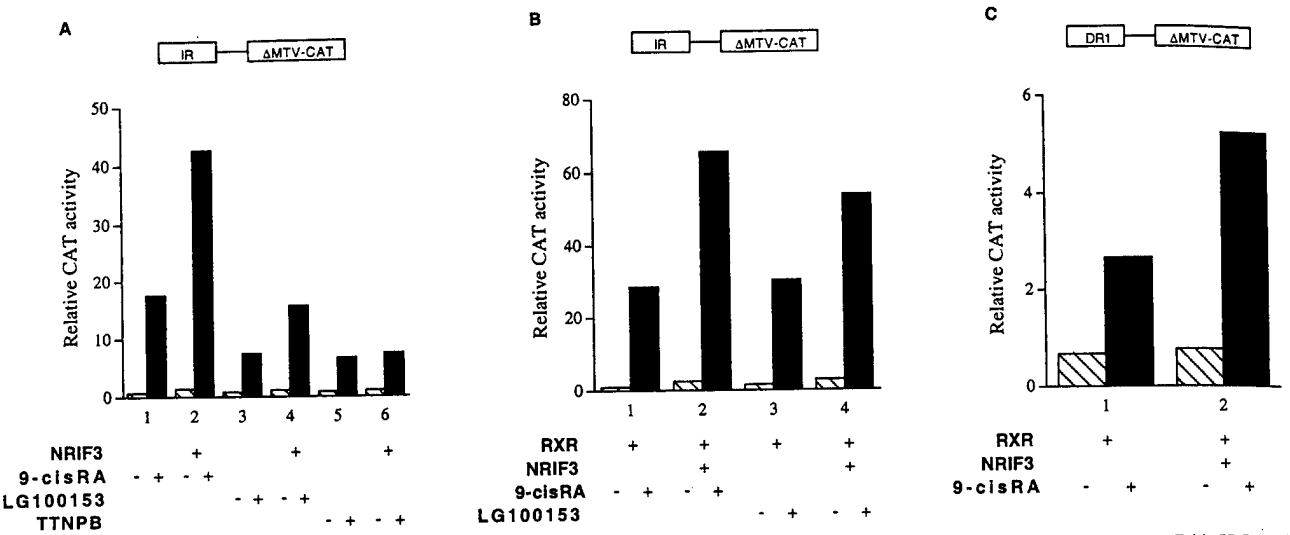


FIG. 6. NRIF3 functions as a coactivator for RXR but not RAR. (A) NRIF3 potentiates the activity of the endogenous RXR(s) but not the RAR(s). HeLa cells were transfected with the IR-ΔMTV-CAT reporter (without any receptor expression vector) to examine the activation by endogenous retinoid receptors. The NRIF3 expression vector or the empty control vector was cotransfected to examine the effect of NRIF3 on the activity of the endogenous RXR(s) or RAR(s). Relative CAT activity was determined in the presence (filled bars) or absence (hatched bars) of the indicated ligands (1 μM). (B and C) NRIF3 potentiates the activity of the exogenously expressed RXR. A vector expressing hRXRα was cotransfected into HeLa cells with the IR-ΔMTV-CAT reporter (B) or the DR1-ΔMTV-CAT reporter (C) in the presence (filled bars) or absence (hatched bars) of the indicated ligands (1 μM). The effect of NRIF3 on RXR-mediated transactivation was examined as described for panel A. TTNPB, a synthetic ligand for RAR.

Cotransfection of NRIF3 enhanced the activation of this reporter by either 9-*cis* RA, or LG100153 (72), an RXR-specific ligand (Fig. 6A, bars 1 and 2 and bars 3 and 4). In contrast, although the RAR-specific ligand TTNPB (68) also activated the IR-ΔMTV-CAT reporter, cotransfection of NRIF3 had no effect (Fig. 6A, bars 5 and 6). These results indicate that NRIF3 potentiates the activity of the endogenous RXR(s) but not the RAR(s), which is consistent with the distinct receptor specificity of NRIF3 revealed from the yeast two-hybrid assay (Table 1) and in vitro binding experiments (Fig. 4).

To further document that NRIF3 can function as a coactivator for RXR, a vector expressing exogenous RXR was cotransfected with IR-ΔMTV-CAT. Exogenous RXR expression enhanced the activation of this CAT reporter by either 9-*cis* RA or LG100153 (compare Fig. 6A and B, bars 1 and 3). This RXR-mediated activation of reporter expression was further stimulated by NRIF3 (Fig. 6B). Finally, we also examined the activation of a DR1-ΔMTV-CAT reporter. This DR1 (AGGT CANAGGTCA [where N is any nucleotide]) sequence is thought to be a specific response element for RXR (39, 51). Although we found that this DR1 is a weaker response element than the IR sequence, cotransfection of an RXR expression vector led to ligand-induced activation of this DR1 reporter, which was also further enhanced by NRIF3 (Fig. 6C).

NRIF3 does not potentiate the activities of GR, PR, ER, and VDR in vivo. The selective coactivation of TR and RXR (but not RAR) by NRIF3 is consistent with its distinct binding specificities to these receptors. To further establish that NRIF3 acts as a receptor-specific coactivator, we next examined the effect of NRIF3 on the activities of four additional nuclear receptors, including GR, PR, ER, and VDR, by transfection experiments. HeLa cells were transfected with a GRE/PRE-tk-CAT reporter along with a vector expressing either GR or PR. As shown in Fig. 7A, cognate hormone treatment results in activation of the CAT reporter. However, expression of NRIF3 has little effect (Fig. 7A). Similar experiments were carried out with ER and ERE-ΔMTV-CAT or VDR and VDRE-ΔMTV-CAT. As shown in Fig. 7B and C, NRIF3 was found to have little or no effect on the activities of these receptors as well. Taken together, the combined results of our transfection studies support the notion that NRIF3 is a coactivator with a unique receptor specificity.

A novel C-terminal domain in NRIF3 is essential for ligand-dependent interactions with TR and RXR. The LXXLL signature motif has been found to be present in the receptor-interacting domains of many identified coactivators, such as SRC-1/NCoA-1 and GRIP1/TIF-2 (32). The broad spectrum of receptor binding by coactivators such as SRC-1 suggests that the LXXLL-containing interacting domain may recognize structurally similar surfaces of these LBDs. Indeed, recent structural and functional studies revealed that the LXXLL motif and its nearby flanking amino acids are involved in direct contact with a hydrophobic cleft of the target surfaces presented by the ligand-bound LBDs of nuclear receptors (19, 23, 52, 56). The facts that NRIF3 also contains an LXXLL motif (amino acids 9 to 13) (Fig. 2 and 8A) and exhibits a distinct receptor specificity raise the possibility that (i) the motif and surrounding amino acids are involved in mediating receptor-specific interaction of NRIF3 or (ii) another region of NRIF3 (alone or in concert with the LXXLL motif region) plays an important role in mediating such an interaction.

To explore these issues, we examined whether EnS and EnL, which contain the same LXXLL motif and flanking amino acids as NRIF3, can interact with nuclear receptors in a yeast two-hybrid assay (Fig. 8). EnS consists of 111 amino acids and is 100% identical to the first 111 residues of NRIF3, while EnL consists of 170 amino acids, the first 161 of which are also 100% identical to the same region in NRIF3 (Fig. 2 legend and Fig. 8A). Thus, NRIF3 and EnL differ only in their C termini, with NRIF3 having a unique region of 16 amino acids and EnL having a unique region of 9 amino acids (Fig. 8A). Interestingly, despite their extensive identity with NRIF3, the interaction with liganded TR or RXR is completely abolished in EnS and EnL (Fig. 8B). We also examined other nuclear receptors that do not interact with NRIF3 and found that they also do not interact with EnS or EnL (data not shown). These results

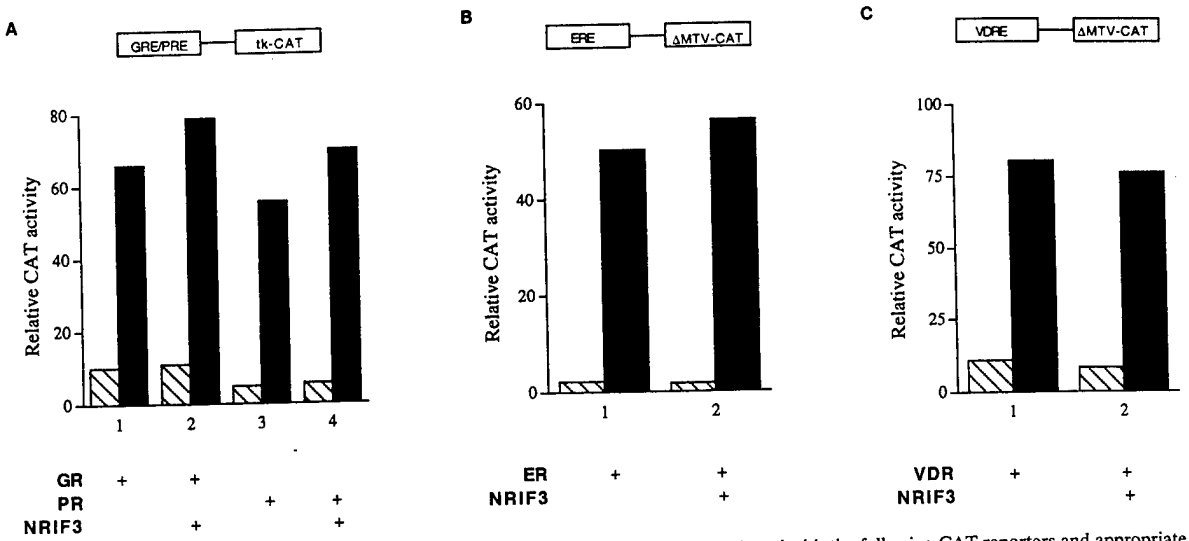


FIG. 7. NRIF3 does not potentiate the activity of GR, PR, ER, or VDR. HeLa cells were transfected with the following CAT reporters and appropriate receptor expression vectors: GRE/PRE-tk-CAT and rGR or hPR (A), ERE-ΔMTV-CAT and hER (B), and VDRE-ΔMTV-CAT and hVDR (C). Cells were incubated in the presence (filled bars) or absence (hatched bars) of 100 nM dexamethasone for GR, progesterone for PR, estradiol for ER, and 1,25-$(OH)_2$-vitamin D_3 for VDR. Cotransfection of NRIF3 was found to have little effect on the activities of these receptors.

indicate that the unique C-terminal domain in NRIF3 (NCD) (residues 162 to 177) is essential for its specific interaction with liganded TR and RXR while the N-terminal LXXLL motif (amino acids 9 to 13) and its flanking sequences are not sufficient to allow for detectable receptor interactions.

Although the LXXLL motif was found to be insufficient for interaction, we examined whether this N-terminal motif of NRIF3 contributes to the NRIF3-receptor interaction by mutating the first leucine of the LXXLL motif into alanine (L9A) by site-directed mutagenesis. Previous experiments have shown that the three leucine residues are essential for an LXXLL module to interact with receptor LBDs and that the replacement of any of them with alanine abolishes the interaction (32). We examined the L9A NRIF3 mutant form for its interaction with TR and RXR in a yeast two-hybrid assay. As shown in Fig. 9, the L9A mutant was still capable of ligand-dependent interaction with TR and RXR (~25-fold induction by ligand). However, the introduced mutation reduced the interaction by about 4-fold (for TR) or 14-fold (for RXR). These results suggest that although the LXXLL motif is not absolutely essential for NRIF3 interaction with liganded receptors, it plays a role in allowing an optimum interaction to occur.

Computer modeling suggests that the NCD docks into the hydrophobic cleft of the liganded LBDs. Secondary-structure analysis of the C-terminal domain of NRIF3 predicted the formation of an α-helix. Moreover, inspection of the putative C-terminal helix revealed an LXXIL motif (amino acids 172 to 176), which is reminiscent of the canonical LXXLL. Although the ultimate elucidation of the molecular basis of the NRIF3-

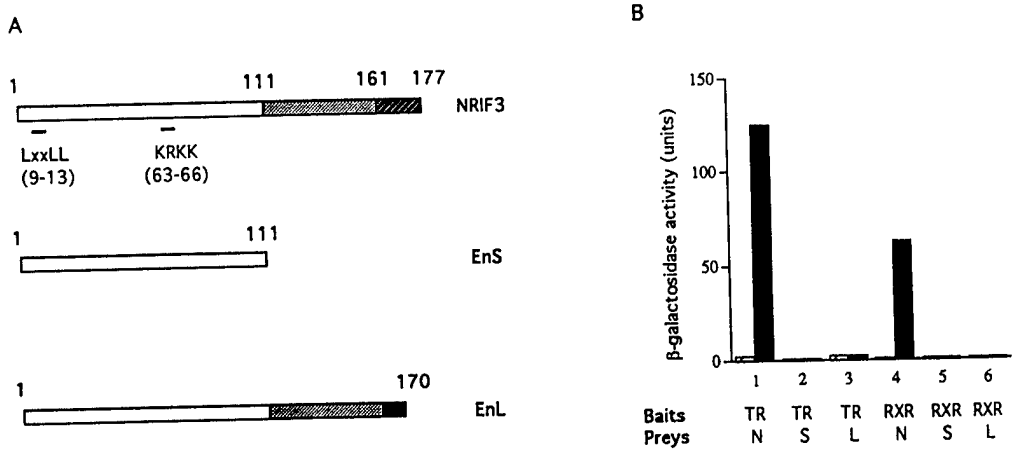


FIG. 8. The NCD is essential for the interaction with liganded TR or RXR. (A) Schematic comparison of NRIF3 with EnS and EnL. EnS is 100% identical to the first 111 amino acids of NRIF3 and EnL (open boxes). The regions from amino acids 112 to 161 in NRIF3 and EnL (stippled boxes) are 100% identical. NRIF3 and EnL differ in their C termini (16 amino acids in NRIF3 [hatched box] and 9 amino acids in EnL [filled box]). The positions of the LXXLL motif and a putative nuclear localization signal (KRKK) are also indicated. (B) NRIF3 (N), EnS (S), or EnL (L) was examined for interaction with LexA-TR or LexA-RXR in a yeast two-hybrid assay as described in Materials and Methods. The assays were performed in the absence (hatched bars) or the presence (filled bars) of 1 μM T3 (for TR) or 9-*cis* RA (for RXR).

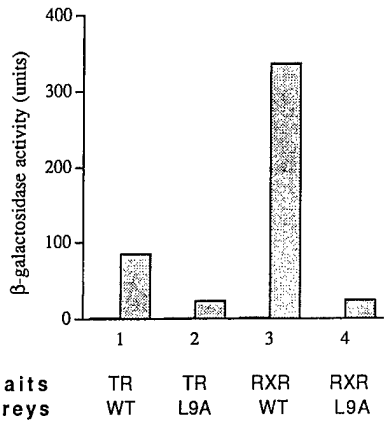


FIG. 9. The LXXLL motif of NRIF3 is required for optimum interaction with TR and RXR. Wild-type NRIF3 (WT) or the L9A NRIF3 mutant (L9A) was examined for interaction with LexA-TR or LexA-RXR in a yeast two-hybrid assay as described in Materials and Methods. β-Galactosidase activities were determined in the absence (filled bars) or presence (stippled bars) of cognate ligands (1 μM T3 for TR; 1 μM 9-*cis* RA for RXR).

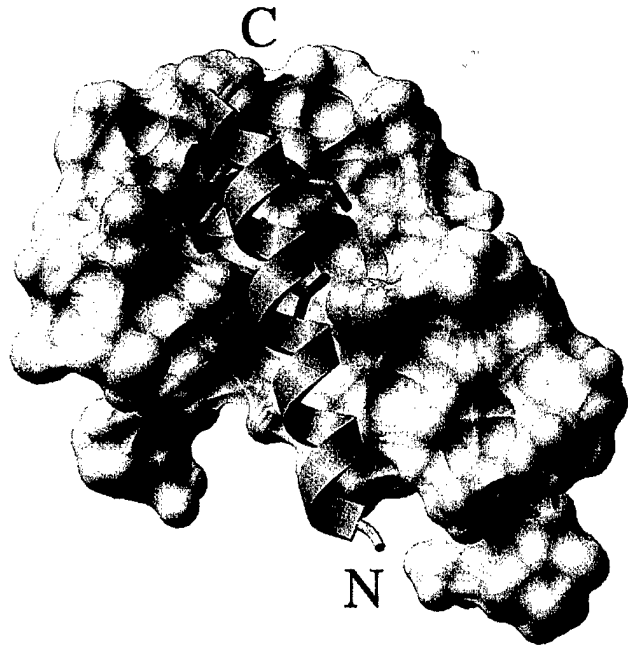


FIG. 10. Hypothetical model of the interaction of the NCD and the liganded LBD. The docking of the C-terminal helix of NRIF3, which contains an LXXIL module, to the ligand-bound LBDs was carried out as described in Materials and Methods. The NCD-TR LBD model is shown here as an example. The side chains of the two leucines (green) and one isoleucine (cyan) of the LXXIL core fit within a hydrophobic groove (salmon) on the surface of the liganded LBD (80). A similar modeling procedure was carried out with an LXXLL box of SRC-1 (result not shown). Putative binding energies (−21 kcal/mol for the NCD and −18 kcal/mol for the LXXLL box of SRC-1) were calculated as described in Materials and Methods. See the text for details.

receptor interaction awaits future studies such as X-ray crystallography, the putative helix structure of the NCD and its LXXIL motif suggest that the NCD may interact with the liganded LBDs in a fashion similar to that of the receptor-interacting domains that employ the canonical LXXLL motif(s). To explore this possibility, we modeled the interaction of the C terminus of NRIF3 with the liganded LBDs, using algorithms developed mainly by the staff of the laboratory of one of the authors (R. Abagyan and coworkers) (1, 63, 70, 74, 75). The background information and procedures used for constructing these models are described in Materials and Methods. The results of our modeling suggest that the NCD fits well into the hydrophobic cleft formed on the LBDs as a result of ligand binding. An example of such a model (NCD-TR LBD) is shown in Fig. 10. In this model, the two leucines and one isoleucine of the LXXIL motif are predicted to be deeply buried in the central cavity of the hydrophobic groove formed by the liganded LBD of the receptor. We also calculated the putative binding energy for the modeled NCD-TR complex, using an improved partitioning binding energy function, with continuum representation of the electrostatics of the system (64). The calculated binding energy for the modeled NCD-TR complex is about −21 kcal/mol. As a control, we carried out a similar modeling procedure using the second LXXLL box within the receptor-interacting domain of SRC-1. This LXXLL box has been shown to be required for interaction with TR (52). Our calculated binding energy for this LXXLL box with liganded TR LBD is −18 kcal/mol, a value that is very close to the one calculated for the NCD. Altogether, our modeling and calculations suggest a mechanism in which the NCD directly mediates interaction with liganded LBDs through an LXXIL motif.

Functional interaction of the NCD with liganded LBDs and the essential role of its LXXIL motif. To explore the possibility suggested from our computer modeling, the NCD (amino acids 162 to 177) was fused to the LexA DNA binding domain and was examined for interaction with the receptor LBDs in a yeast two-hybrid assay. The LexA-NCD fusion protein alone does not activate the LacZ reporter in yeast (data not shown). As a negative control, we also found that LexA-NCD does not interact with the B42 activation domain itself (Fig. 11) and that LexA alone does not interact with the receptor LBDs (data not shown). However, when the LexA-NCD and the LBD of TR or RXR (fused with B42) were used in the two-hybrid assay, a strong ligand-dependent interaction was observed, as indicated by the induction of β-galactosidase activity by their cognate

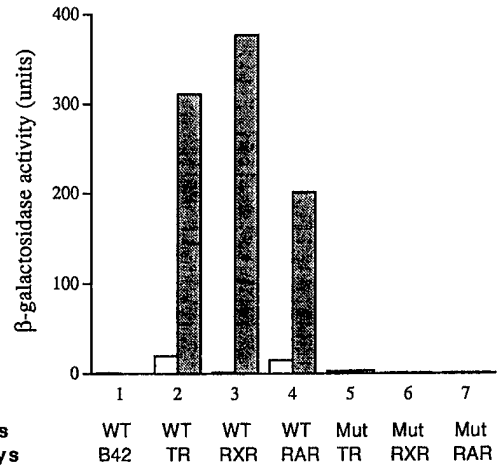


FIG. 11. Interaction of the NCD with the receptor LBDs and the role of the LXXIL motif. The wild-type NCD (WT) or the NCD mutant form (Mut) in which the three core hydrophobic residues of the LXXIL motif (two leucines and one isoleucine) are changed into alanines was examined for interaction with the LBDs of TR, RXR, and RAR in a yeast two-hybrid assay as described in Materials and Methods. β-Galactosidase activities were determined in the absence (open bars) or presence (stippled bars) of cognate ligands. The prey expressing B42 alone was used as a negative control.

ligands (Fig. 11). These results suggest that the NCD can directly interact with the LBDs of TR and RXR in a ligand-dependent manner.

Since NRIF3 harbors a distinct receptor specificity in interacting only with TR and RXR and not other receptors (e.g., RAR), we next asked whether the NCD also harbors a receptor specificity. To our surprise, the NCD was found to interact efficiently with the LBD of RAR in a ligand-dependent manner (Fig. 11). Therefore, while our results clearly suggest that the NCD is an important surface for receptor interactions, as the NCD is found to be both essential for (Fig. 8) and sufficient to mediate (Fig. 11) such interactions, it nevertheless does not appear to be (solely) responsible for the receptor specificity of NRIF3. It is possible that another region of the NRIF3 molecule contributes to the observed receptor specificity of NRIF3 and/or that the specificity is determined by the overall three-dimensional structure of NRIF3.

Since our model predicts the importance of the LXXIL motif in the NCD-receptor interaction (Fig. 10), we tested this by changing the three core residues of the motif (two leucines and one isoleucine) into alanines. As expected, interaction with the LBDs is completely abolished in the resulting mutant NCD (Fig. 11), confirming that the LXXIL motif is essential for the interaction.

DISCUSSION

Recent efforts in understanding receptor-mediated transcription have led to the identification of a number of coactivators for nuclear hormone receptors, which can be categorized into two main groups based on overall homology, the SRC-1 family (including SRC-1/NCoA-1, TIF2/GRIP1/NCoA-2, and AIB1/p/CIP/ACTR/RAC3/TRAM-1) (2, 14, 34, 35, 37, 44, 58, 73, 74, 79) and the CBP/p300 family (13, 31, 37). Other putative coactivators (e.g., ARA70 and PGC-1) that do not belong to the SRC-1 or CBP/p300 family have also been identified (60, 85). In addition, p/CAF may also be involved in receptor action through its association with nuclear receptors as well as with other coactivators (11, 14, 38, 83). Among these known coactivators, CBP/p300, members of the SRC-1 group, and p/CAF all possess histone acetyltransferase activities (8, 14, 57, 69, 83).

In this study we report the identification of a novel nuclear protein (NRIF3) which exhibits specific ligand-dependent interactions with TR and RXR but not with RAR, VDR, GR, PR, or ER. Functional experiments indicated that NRIF3 potentiates TR- and RXR-mediated transactivation in vivo but exhibits little or no effect on the activities of other examined receptors. Therefore, NRIF3 represents a novel coactivator with a distinct receptor specificity and, thus, may shed light on clarifying the molecular mechanism(s) underlying receptor-specific regulation of gene expression.

A database search indicated that NRIF3 has no homology with any known coactivators except in a single LXXLL motif. An unusual feature of NRIF3 is its relatively small size, which is in sharp contrast to the sizes of SRC-1 and CBP/p300. A homology search identified two alternatively spliced isoforms of NRIF3 which were previously designated β3-endonexin short and long forms (67). Preliminary studies with these two endonexins indicate that, like NRIF3, they localize to the cell nucleus (43a, 66a). Interestingly, despite their extensive identities with NRIF3, both EnS and EnL fail to exhibit interaction with liganded nuclear receptors (Fig. 8). Consistent with this finding, we found that EnS and EnL have little effect on receptor-mediated transcription in transfection experiments (data not shown). Therefore, the precise roles of these two endonexins remain to be elucidated. We suggest two not mutually exclusive possibilities. First, since both EnL and EnS appear to localize to the nucleus, it is possible that they act as cofactors for other transcriptional regulators. Second, since the EnS can interact with the cytoplasmic tail of β3-integrin (22, 67), it may communicate signals generated at the plasma membrane to the cell nucleus. An example of a protein which is involved in both cell adhesion and transcriptional regulation is β-catenin (82).

Previous study of the endonexins identified the presence of NRIF3-related mRNAs (by Northern blotting) in a wide range of human tissues (67). Because NRIF3 and EnL contain almost identical nucleotide sequences and differ only by an alternative splice site which results in the removal of a small exon in NRIF3, it is difficult to specifically identify NRIF3 mRNA by Northern blotting. A search of the expressed sequence tag database indicates that NRIF3 mRNA, as well as both EnL and EnS mRNAs, is expressed. However, the precise determination of cell and tissue distribution of NRIF3, EnS, and EnL will require the development of highly selective antibodies. Nevertheless, the wide expression pattern of NRIF3-related mRNAs is consistent with the role of NRIF3 as a coactivator of the TRs, which are also widely expressed (70), or the RXRs, which are ubiquitously expressed (48).

A key goal concerning the action of nuclear hormone receptors is to understand the molecular events underlying the functional specificities of different receptors in regulating the expression of their target genes. Determinants of specificity include specific ligand binding and selective binding of the receptors to their cognate response elements, as well as specific expression patterns of different receptors. These determinants alone, however, are not always sufficient to explain the extents of specificity observed for members of the nuclear receptor family. For example, several members of the thyroid hormone/retinoid receptor subfamily may bind similarly to common DNA elements while target genes containing those elements are only selectively activated by certain receptors (20, 47). Therefore, it is likely that additional factors (determined by cell and promoter contexts) are involved in determining receptor functional specificity. In this respect, most known coactivators do not appear to be receptor specific. For example, members of the SRC-1 and CBP/p300 families interact with and appear to be involved in the actions of many nuclear receptors (13, 14, 34, 37). Two known coactivators that may be involved in receptor-specific functions are ARA70 and PGC-1. The AR coactivator ARA70 has been reported to potentiate the activity of AR more efficiently than it does the activities of other nuclear receptors (85). However, whether ARA70 can associate with other receptors remains to be thoroughly examined. The expression of PGC-1 is restricted mainly to the brown fat tissue and is thought to be directly involved in the regulation of thermogenesis by PPARγ (60). Nevertheless, PGC-1 exhibits a relatively broad spectrum of binding to different nuclear receptors. Therefore, the identification of NRIF3 represents the first example of a coactivator with such a clearly defined receptor specificity.

The receptor specificity of NRIF3 raises an interesting question about its molecular mechanism. Domain analysis suggests that the LXXLL motif (amino acids 9 to 13) and its flanking sequences in NRIF3 are not sufficient for interaction with liganded nuclear receptors. In fact, such interaction is completely abolished in EnL, an alternatively spliced product which has the same LXXLL motif and contains the first 161 amino acids (of a total of 177 amino acids) of NRIF3. This result suggests that a putative domain consisting of the last 16 amino acids of NRIF3 (residues 162 to 177) is essential for its

interaction with liganded receptors. Inspection of this NCD indicates that it contains an LXXIL motif (amino acids 172 to 176), and secondary-structure analysis predicts the formation of an α-helix. The predicted helix structure and the similarity of LXXIL to the canonical LXXLL raise the possibility that this LXXIL-containing region plays a direct role in NRIF3-receptor interactions.

Our modeling of the NCD-LBD interaction (Fig. 10) suggests that the same hydrophobic groove in the ligand-bound LBD, which has been shown by previous studies to be the binding site for coactivators such as SRC-1/NCoA-1 and GRIP1 (19, 23, 56), may also be a suitable site for the docking of the C-terminal helix of NRIF3. Thus, the utilization of the complementary pair of an α-helix (in the coactivator) and a hydrophobic groove (in the receptor) for interaction seems to be a general scheme that also applies to NRIF3. The binding energy estimated for the NCD and the TR LBD (−21 kcal/mol) is similar to the value calculated for the second LXXLL box of SRC-1/NCoA-1 and the TR LBD (−18 kcal/mol). To explore the mechanisms suggested by the modeling, we found that the NCD can directly mediate interaction with the LBDs in a ligand-dependent manner (Fig. 11). Moreover, the LXXIL motif contained in the NCD was found to be essential for such interactions (Fig. 11). In summary, the results of a combination of a computer modeling and domain and mutagenesis analyses clearly suggest that the NCD is an important surface that is directly involved in interaction with the LBDs of the receptors, where the LXXIL motif of the NCD mimics the function of a canonical LXXLL. The AF-2 helix (which is a critical constituent of the hydrophobic groove formed upon ligand binding) of the LBD has been shown to be important for interaction with the LXXLL boxes of the coactivators (23). Interestingly, we have examined two TR AF-2 mutants (66) and found that in both cases, ligand-dependent interaction with NRIF3 was abolished (43a).

However, the NCD alone does not appear to harbor the same specificity as NRIF3 (Fig. 11). Thus, it seems likely that another part of the NRIF3 molecule contributes to the observed specificity and/or that the specificity is determined by the overall three-dimensional structure of NRIF3. In this regard, the potential role of the N-terminal LXXLL motif is intriguing. Although the N-terminal LXXLL motif (amino acids 9 to 13) is insufficient alone to mediate an interaction with TR or RXR (Fig. 8), it can influence the interaction of NRIF3 with these receptors, as the L9A NRIF3 mutant retains a significant but nevertheless reduced level of association with liganded TR or RXR (Fig. 9). Thus, NRIF3 appears to employ at least two regions in interacting with liganded TR or RXR, with the NCD playing an essential role and the N-terminal LXXLL motif playing a secondary role. A simplified explanation for these findings is that the NCD provides a major surface for receptor binding while the N-terminal LXXLL motif makes some minor contact with either the same receptor molecule or, more likely, with the other partner of a homodimer or heterodimer to further stabilize the NRIF3-receptor interaction. An example of a coactivator molecule employing two separate regions to interact with the two partners of a receptor dimer has been documented for the recently solved crystal structure of liganded PPARγ complexed with SRC-1/NCoA-1 (56). If NRIF3 indeed employs both its NCD and its N-terminal LXXLL motif in receptor interactions, the specificity may result from the intramolecular dialog between the two regions as well as the intermolecular dialog among NRIF3 and the receptors. However, it remains possible that the N-terminal LXXLL plays only a more indirect role and that the overall three-dimensional structure of NRIF3 is responsible for its observed specificity.

Accumulating evidence suggests that the actions of transcriptional activating proteins are (usually) mediated by multiprotein complexes (59), and such a scheme is also likely for nuclear receptors. For example, biochemical evidence suggests that multiprotein complexes associate with liganded TR and VDR (24, 62, 86). Interestingly, many of the proteins identified in these studies are not known coactivators. While the study of known coactivators such as CBP/p300 and members of the SRC-1 family has suggested that histone acetylation may play an important role in receptor-mediated transactivation (8, 14, 57, 69), detailed elucidation of the transactivation mechanism(s) used by these receptors awaits the identification and study of additional cofactors involved in the transactivation process.

Our results with NRIF3 suggest that transcriptional activation by nuclear receptors may employ a receptor-specific coactivator(s) in addition to the generally used coactivators such as CBP and SRC-1. Therefore, coactivators with NRIF3-like properties may contribute to the functional specificities of nuclear receptors in vivo. Based on our results with NRIF3 and the results of previous studies of nuclear receptor action, we suggest a combinatorial specificity model where a coactivation complex is likely composed of two kinds of factors: general factors that interact with and are involved in the action of many nuclear receptors (such as CBP and SRC-1) and specific factors that exhibit receptor specificity (such as NRIF3). In addition to interacting with the liganded receptor, coactivators may also communicate with each other within the coactivation complex through protein-protein interactions (e.g., CBP/p300 can interact with SRC-1/NCoA-1 or p/CIP) (37, 74, 84). An intriguing possibility is that the combinatorial actions of specific factors and other partners involved in the transactivation process facilitate the recruitment of specific coactivation complexes for different receptors (under different cell, promoter, and HRE contexts), which would provide an important mechanistic layer for receptor functional specificity. An advantage of employing such a combinatorial strategy is that a broad array of diversity can be generated from a relatively small number of involved factors. Further study of NRIF3 with known and possibly other yet to be identified coactivators, as well as analysis of the interplay among these coactivators, should provide important insights into the molecular mechanism(s) underlying the specificity of receptor-mediated regulation of target gene expression.

ACKNOWLEDGMENTS

We are grateful to Richard Goodman, Bert O'Malley, Ming-Jer Tsai, Michael Garabedian, J. Wesley Pike, Gunther Schutz, Ron Evans, and David Moore for plasmids and Richard Heyman of Ligand, Inc., for providing the retinoids. We thank Sanford Shattil for GFP-EnL and GFP-EnS; Gordon Fishell for help with fluorescence microscopy; Chun Wong, Ula Huang, Sidney Guo, and Paul Modlinger for experimental assistance; and Bruce Raaka and Fred Stanley for advice with graphic preparations.

This research was supported by NIH grant DK16636 (to H.H.S.), NRSA postdoctoral fellowship award DK09581 (to D.L.), DOD grant DAMD179818133, NIH grant GM5541801, DOE grant DEFG0296 ER62268 (to R.A. and M.S.), and an NIH short-term training grant for students in health professional schools (DK07421) (to E.L.). V.D.-Y. was supported in part by The Aaron Diamond Foundation (grant HRI817-5332F). H.H.S., V.D.-Y., and R.A. are members of the NYUMC Cancer Center (grant CA16087). Sequence analysis and database searches were through the NYUMC Research Computing Resource, which received support from the National Science Foundation (grant DIR-8908095).

REFERENCES

1. **Abagyan, R. A., M. M. Totrov, and D. A. Kuznetsov.** 1994. ICM: a new method for structure modeling and design: applications to docking and structure prediction from the distorted native conformation. J. Comp. Chem. **15**:488–506.
2. **Anzick, S. L., J. Kononen, R. L. Walker, D. O. Azorsa, M. M. Tanner, X. Y. Guan, G. Sauter, O. P. Kallioniemi, J. M. Trent, and P. S. Meltzer.** 1997. AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. Science **277**:965–968.
3. **Arany, Z., D. Newsome, E. Oldread, D. M. Livingston, and R. Eckner.** 1995. A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. Nature **374**:81–84.
4. **Arias, J., A. S. Alberts, P. Brindle, F. X. Claret, T. Smeal, M. Karin, J. Feramisco, and M. Montminy.** 1994. Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. Nature **370**:226–229.
5. **Au-Fliegner, M., E. Helmer, J. Casanova, B. M. Raaka, and H. H. Samuels.** 1993. The conserved ninth C-terminal heptad in thyroid hormone and retinoic acid receptors mediates diverse responses by affecting heterodimer but not homodimer formation. Mol. Cell. Biol. **13**:5725–5737.
6. **Baniahmad, A., X. Leng, T. P. Burris, S. Y. Tsai, M.-J. Tsai, and B. W. O'Malley.** 1995. The τ4 activation domain of the thyroid hormone receptor is required for release of a putative corepressor(s) necessary for transcriptional silencing. Mol. Cell. Biol. **15**:76–86.
7. **Bannister, A. J., and T. Kouzarides.** 1995. CBP-induced stimulation of c-Fos activity is abrogated by E1A. EMBO J. **14**:4758–4762.
8. **Bannister, A. J., and T. Kouzarides.** 1996. The CBP co-activator is a histone acetyltransferase. Nature **384**:641–643.
9. **Barettino, D., M. D. M. V. Ruiz, and H. G. Stunnenberg.** 1994. Characterization of the ligand-dependent transactivation domain of thyroid hormone receptor. EMBO J. **13**:3039–3049.
10. **Bhattacharya, S., R. Eckner, S. Grossman, E. Oldread, Z. Arany, A. D'Andrea, and D. M. Livingston.** 1996. Cooperation of Stat2 and p300/CBP in signalling induced by interferon-alpha. Nature **383**:344–347.
11. **Blanco, J. C. G., S. Minucci, J. Lu, X. J. Yang, K. K. Walker, H. Chen, R. M. Evans, V. Nakatani, and K. Ozato.** 1998. The histone acetylase PCAF is a nuclear receptor coactivator. Genes Dev. **12**:1638–1651.
12. **Casanova, J., E. Helmer, S. Selmi-Ruby, J.-S. Qi, M. Au-Fliegner, V. Desai-Yajnik, N. Koudinova, F. Yarm, B. M. Raaka, and H. H. Samuels.** 1994. Functional evidence for ligand-dependent dissociation of thyroid hormone and retinoic acid receptors from an inhibitory cellular factor. Mol. Cell. Biol. **14**:5756–5765.
13. **Chakravarti, D., V. J. LaMorte, M. C. Nelson, T. Nakajima, I. G. Schulman, H. Juguilon, M. Montminy, and R. M. Evans.** 1996. Role of CBP/P300 in nuclear receptor signalling. Nature **383**:99–103.
14. **Chen, H., R. J. Lin, R. L. Schiltz, D. Chakravarti, A. Nash, L. Nagy, M. L. Privalsky, Y. Nakatani, and R. M. Evans.** 1997. Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. Cell **90**:569–580.
15. **Chen, J. D., and R. M. Evans.** 1995. A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature **377**:454–457.
16. **Chrivia, J. C., R. P. Kwok, N. Lamb, M. Hagiwara, M. R. Montminy, and R. H. Goodman.** 1993. Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature **365**:855–859.
17. **Conneely, O. M., W. P. Sullivan, D. O. Toft, M. Birnbaumer, R. G. Cook, B. L. Maxwell, T. Zaraucko-Schulz, G. L. Greene, W. T. Schraeder, and B. W. O'Malley.** 1986. Molecular cloning of the chicken progesterone receptor. Science **233**:767–770.
18. **Cormack, B. P., R. H. Valdivia, and S. Falkow.** 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene **173**:33–38.
19. **Darimont, B. D., R. L. Wagner, J. W. Apriletti, M. R. Stallcup, P. J. Kushner, D. Baxter, R. J. Fletterick, and K. R. Yamamoto.** 1998. Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev. **12**:3343–3356.
20. **Desai-Yajnik, V., and H. H. Samuels.** 1993. The NF-κB and Sp1 DNA motifs of the human immunodeficiency virus type 1 long terminal repeat function as novel thyroid hormone response elements. Mol. Cell. Biol. **13**:5057–5069.
21. **Durand, B., M. Saunders, C. Gausdon, B. Roy, R. Losson, and P. Chambon.** 1994. Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. EMBO J. **13**:5370–5382.
22. **Eigenthaler, M., L. Hofferer, S. J. Shattil, and M. H. Ginsberg.** 1997. A conserved sequence motif in the integrin β3 cytoplasmic domain is required for its specific interaction with β3-endonexin. J. Biol. Chem. **272**:7693–7698.
23. **Feng, W., R. C. Ribeiro, R. L. Wagner, H. Nguyen, J. W. Apriletti, R. J. Fletterick, J. D. Baxter, P. J. Kushner, and B. L. West.** 1998. Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. Science **280**:1747–1749.
24. **Fondell, J. D., H. Ge, and R. G. Roeder.** 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. Proc. Natl. Acad. Sci. USA **93**:8329–8333.
25. **Forman, B. M., J. Casanova, B. M. Raaka, J. Ghysdael, and H. H. Samuels.** 1992. Half-site spacing and orientation determines whether thyroid hormone and retinoic acid receptors and related factors bind to DNA response elements as monomers, homodimers, or heterodimers. Mol. Endocrinol. **6**:429–442.
26. **Giguere, V., E. S. Ong, P. Segui, and R. M. Evans.** 1987. Identification of a receptor for the morphogen retinoic acid. Nature **330**:624–629.
27. **Glass, C. K., D. W. Rose, and M. G. Rosenfeld.** 1997. Nuclear receptor coactivators. Curr. Opin. Cell Biol. **9**:222–232.
28. **Gu, W., X. L. Shi, and R. G. Roeder.** 1997. Synergistic activation of transcription by CBP and p53. Nature **387**:819–823.
29. **Gyuris, J., E. Golemis, H. Chertkov, and R. Brent.** 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell **75**:791–803.
30. **Hadzic, E., V. Desai-Yajnik, E. Helmer, S. Guo, S. Wu, N. Koudinova, J. Casanova, B. M. Raaka, and H. H. Samuels.** 1995. A 10-amino-acid sequence in the N-terminal A/B domain of thyroid hormone receptor α is essential for transcriptional activation and interaction with the general transcription factor TFIIB. Mol. Cell. Biol. **15**:4507–4517.
31. **Hanstein, B., R. Eckner, J. DiRenzo, S. Halachmi, H. Liu, B. Searcy, R. Kurokawa, and M. Brown.** 1996. p300 is a component of an estrogen receptor coactivator complex. Proc. Natl. Acad. Sci. USA **93**:11540–11545.
32. **Heery, D. M., E. Kalkhoven, S. Hoare, and M. G. Parker.** 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature **387**:733–736.
33. **Heinzel, T., R. M. Lavinsky, T. M. Mullen, M. Soderstrom, C. D. Laherty, J. Torchia, W. M. Yang, G. Brard, S. D. Ngo, J. R. Davie, E. Seto, R. N. Eisenman, D. W. Rose, C. K. Glass, and M. G. Rosenfeld.** 1997. A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature **387**:43–48.
34. **Hong, H., K. Kohli, M. Garabedian, and M. R. Stallcup.** 1997. GRIP1, a transcriptional coactivator for the AF-2 transactivation domain of steroid, thyroid, retinoid, and vitamin D receptors. Mol. Cell. Biol. **17**:2735–2744.
35. **Hong, H., K. Kohli, A. Trivedi, D. L. Johnson, and M. R. Stallcup.** 1996. GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. Proc. Natl. Acad. Sci. USA **93**:4948–4952.
36. **Horlein, A. J., A. M. Naar, T. Heinzel, J. Torchia, B. Gloss, R. Kurokawa, A. Ryan, Y. Kamil, M. Soderstrom, C. K. Glass, and M. G. Rosenfeld.** 1995. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature **377**:397–404.
37. **Kamei, Y., L. Xu, T. Heinzel, J. Torchia, R. Kurokawa, B. Gloss, S.-C. Lin, R. A. Heyman, D. W. Rose, C. K. Glass, and M. G. Rosenfeld.** 1996. A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell **85**:403–414.
38. **Korzus, E., J. Torchia, D. W. Rose, L. Xu, R. Kurokawa, E. M. McInerney, T. M. Mullen, C. K. Glass, and M. G. Rosenfeld.** 1998. Transcription factor-specific requirements for coactivators and their acetyltransferase functions. Science **279**:703–707.
39. **Kurokawa, R., J. DiRenzo, M. Boehm, J. Sugarman, B. Glass, M. G. Rosenfeld, R. A. Heyman, and C. K. Glass.** 1994. Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. Nature **371**:528–531.
40. **Kwok, R. P., J. R. Lundblad, J. C. Chrivia, J. P. Richards, H. P. Bachinger, R. G. Brennan, S. G. Roberts, M. R. Green, and R. H. Goodman.** 1994. Nuclear protein CBP is a coactivator for the transcription factor CREB. Nature **370**:223–226.
41. **Lanz, R. B., N. J. McKenna, S. A. Onate, U. Albrecht, J. Wong, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley.** 1999. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell **97**:17–27.
42. **Lee, J. W., H.-S. Choi, J. Gyuris, R. Brent, and D. D. Moore.** 1995. Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. Mol. Endocrinol. **9**:243–254.
43. **Leng, X., J. Blanco, S. Y. Tsai, K. Ozato, B. W. O'Malley, and M.-J. Tsai.** 1995. Mouse retinoid X receptor contains a separable ligand-binding and transactivation domain in its E region. Mol. Cell. Biol. **15**:255–263.

43a. **Li, D., and H. H. Samuels.** Unpublished data.

44. **Li, H., P. J. Gomes, and J. D. Chen.** 1997. RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. Proc. Natl. Acad. Sci. USA **94**:8479–8484.
45. **Lill, N. L., S. R. Grossman, D. Ginsberg, J. DeCaprio, and D. M. Livingston.** 1997. Binding and modulation of p53 by p300/CBP coactivators. Nature **387**:823–827.
46. **Lundblad, J. R., R. P. Kwok, M. E. Laurance, M. L. Harter, and R. H. Goodman.** 1995. Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. Nature **374**:85–88.
47. **MacDonald, P. N., D. R. Dowd, S. Nakajima, M. A. Galligan, M. C. Reeder, C. A. Haussler, K. Ozato, and M. R. Haussler.** 1993. Retinoid X receptors stimulate and 9-*cis* retinoic acid inhibits 1,25-dihydroxyvitamin D_3-activated expression of the rat osteocalcin gene. Mol. Cell. Biol. **13**:5907–5917.
48. **Mangelsdorf, D. J., U. Borgmeyer, R. A. Heyman, J. Y. Zhou, E. S. Ong, A. E. Oro, A. Kakizuka, and R. M. Evans.** 1992. Characterization of three RXR genes that mediate the action of 9-*cis* retinoic acid. Genes Dev. **6**:329–344.

49. **Mangelsdorf, D. J., and R. M. Evans.** 1995. The RXR heterodimers and orphan receptors. Cell **83**:841–850.
50. **Mangelsdorf, D. J., E. S. Ong, J. A. Dyck, and R. M. Evans.** 1990. Nuclear receptor that identifies a novel retinoic acid response pathway. Nature **345**: 224–229.
51. **Mangelsdorf, D. J., C. Thummel, M. Beato, P. Herrlich, G. Schutz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon, and R. M. Evans.** 1995. The nuclear receptor superfamily: the second decade. Cell **83**: 835–839.
52. **McInerney, E. M., D. W. Rose, S. E. Flynn, S. Westin, T. M. Mullen, A. Krones, J. Inostroza, J. Torchia, R. T. Nolte, N. Assa-Munt, M. V. Milburn, C. K. Glass, and M. G. Rosenfeld.** 1998. Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation. Genes Dev. **12**:3357–3368.
53. **Miesfeld, R., S. Rusconi, P. J. Godowski, B. A. Maler, S. Okret, A.-C. Wikstrom, J.-A. Gustafsson, and K. R. Yamamoto.** 1986. Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. Cell **46**:389–399.
54. **Nagpal, S., S. Friant, H. Nakshatri, and P. Chambon.** 1993. RARs and RXRs: evidence for two autonomous transactivation functions (AF-1 and AF-2) and heterodimerization *in vivo*. EMBO J. **12**:2349–2360.
55. **Nagy, L., H. V. Kao, D. Chakravarti, R. J. Lin, C. A. Hassig, D. E. Ayer, S. L. Schreiber, and R. M. Evans.** 1997. Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell **89**: 373–380.
56. **Nolte, R. T., G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass, and M. V. Milburn.** 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. Nature **395**:137–143.
57. **Ogryzko, V. V., R. L. Schiltz, V. Russanova, B. H. Howard, and Y. Nakatani.** 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell **87**:953–959.
58. **Onate, S. A., S. Y. Tsai, M.-J. Tsai, and B. W. O'Malley.** 1995. Sequence and characterization of a coactivator of the steroid hormone receptor superfamily. Science **270**:1354–1357.
59. **Orphanides, G., T. Lagrange, and D. Reinberg.** 1996. The general transcription factors of RNA polymerase II. Genes Dev. **10**:2657–2683.
60. **Puigserver, P., Z. Wu, C. W. Park, R. Graves, M. Wright, and B. M. Spiegelman.** 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. Cell **92**:829–839.
61. **Qi, J.-S., V. Desai-Yajnik, M. E. Greene, B. M. Raaka, and H. H. Samuels.** 1995. The ligand binding domains of the thyroid hormone/retinoid receptor gene subfamily function in vivo to mediate heterodimerization, gene silencing, and transactivation. Mol. Cell. Biol. **15**:1817–1825.
62. **Rachez, C., Z. Suldan, J. Ward, C. P. Chang, D. Burakov, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman.** 1998. A novel protein complex that interacts with the vitamin D3 receptor in a ligand-dependent manner and enhances VDR transactivation in a cell-free system. Genes Dev. **12**: 1787–1800.
63. **Renaud, J.-P., N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer, and D. Moras.** 1995. Crystal structure of the RAR-γ ligand-binding domain bound to all-*trans* retinoic acid. Nature **378**:681–689.
64. **Schapira, M., M. Totrov, and R. A. Abagyan.** 1999. Prediction of the binding energy for small molecules, peptides and proteins. J. Mol. Recognit. **12**: 177–190.
65. **Schule, R., K. Umesono, D. J. Mangelsdorf, J. Bolado, J. W. Pike, and R. M. Evans.** 1990. Jun-Fos and receptors for vitamins A and D recognize a common response element in the human osteocalcin gene. Cell **61**:497–504.
66. **Selmi-Ruby, S., J. Casanova, S. Malhotra, B. Roussett, B. M. Raaka, and H. H. Samuels.** 1998. Role of the conserved C-terminal region of thyroid hormone receptor-α in ligand-dependent transcriptional activation. Mol. Cell. Endocrinol. **138**:105–114.

66a. **Shattil, S. J.** Personal communication.

67. **Shattil, S. J., T. O'Toole, M. Eigenthaler, V. Thon, M. Williams, B. M. Babior, and M. H. Ginsberg.** 1995. β3-Endonexin, a novel polypeptide that interacts specifically with the cytoplasmic tail of the integrin β3 subunit. J. Cell Biol. **131**:807–816.
68. **Sheikh, M. S., Z.-M. Shao, X.-S. Li, M. Dawson, A. M. Jetten, S. Wu, B. A. Conley, M. Garcia, H. Rochefort, and J. A. Fontana.** 1994. Retinoid-resistant estrogen receptor-negative human breast carcinoma cells transfected with retinoic acid receptor-α acquire sensitivity to growth inhibition by retinoids. J. Biol. Chem. **269**:21440–21447.
69. **Spencer, T. E., G. Jenster, M. M. Burcin, C. D. Allis, J. Zhou, C. A. Mizzen, N. J. McKenna, S. A. Onate, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley.** 1997. Steroid receptor coactivator-1 is a histone acetyltransferase. Nature **389**: 194–198.
70. **Strait, K., H. L. Schwartz, A. Perez-Castillo, and J. H. Oppenheimer.** 1990. Relationship of c-erbA mRNA content to tissue triiodothyronine nuclear binding capacity and function in developing and adult rats. J. Biol. Chem. **265**:10514–10521.
71. **Strynadka, N. C., M. Eisenstein, E. Katchalski-Katzir, B. K. Shoichet, I. D. Kuntz, R. Abagyan, M. Totrov, J. Janin, J. Cherfils, F. Zimmerman, A. Olson, B. Duncan, M. Rao, R. Jackson, M. Sternberg, and M. N. James.** 1996. Molecular docking programs successfully predict the binding of a beta-lactamase inhibitory protein to TEM-1 beta-lactamase. Nat. Struct. Biol. **3**:233–239.
72. **Sugawara, A., N. Sanno, N. Takahashi, R. Y. Osamura, and K. Abe.** 1997. Retinoid X receptors in the kidney: their protein expression and functional significance. Endocrinology **138**:3175–3180.
73. **Takeshita, A., G. R. Cardona, N. Koibuchi, C. S. Suen, and W. W. Chin.** 1997. TRAM-1, a novel 160-kDa thyroid hormone receptor activator molecule, exhibits distinct properties from steroid receptor coactivator-1. J. Biol. Chem. **272**:27629–27634.
74. **Torchia, J., D. W. Rose, J. Inostroza, Y. Kamei, S. Westin, C. K. Glass, and M. G. Rosenfeld.** 1997. The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature **387**:677–684.
75. **Totrov, M., and R. Abagyan.** 1994. Detailed ab initio prediction of lysozyme-antibody complex with 1.6 Å accuracy. Nat. Struct. Biol. **1**:259–263.
76. **Totrov, M., and R. Abagyan.** 1997. Flexible protein-ligand docking by global energy optimization in internal coordinates. Proteins **1997** (Suppl. 1):215–220.
77. **Umesono, K., and R. M. Evans.** 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell **57**:1139–1146.
78. **Umesono, K., K. K. Murakami, C. C. Thompson, and R. M. Evans.** 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D receptors. Cell **65**:1255–1266.
79. **Voegel, J. J., M. J. S. Heine, C. Zechel, P. Chambon, and H. Gronemeyer.** 1996. TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J. **15**:3667–3675.
80. **Wagner, R. L., J. W. Apriletti, M. E. McGrath, B. L. West, J. D. Baxter, and R. J. Fletterick.** 1995. A structural role for hormone in the thyroid hormone receptor. Nature **378**:690–697.
81. **Waterman, M. L., S. Adler, C. Nelson, G. L. Greene, R. M. Evans, and M. G. Rosenfeld.** 1988. A single domain of the estrogen receptor confers deoxyribonucleic acid binding and transcriptional activation of the rat prolactin gene. Mol. Endocrinol. **2**:14–21.
82. **Willert, K., and R. Nusse.** 1998. Beta-catenin: a key mediator of Wnt signaling. Curr. Opin. Genet. Dev. **8**:95–102.
83. **Yang, X. J., V. V. Ogryzko, J. Nishikawa, B. H. Howard, and Y. Nakatani.** 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature **382**:319–324.
84. **Yao, T. P., G. Ku, N. Zhou, R. Scully, and D. M. Livingston.** 1996. The nuclear hormone receptor coactivator SRC-1 is a specific target of p300. Proc. Natl. Acad. Sci. USA **93**:10626–10631.
85. **Yeh, S., and C. Chang.** 1996. Cloning and characterization of a specific coactivator, ARA70, for the androgen receptor in human prostate cells. Proc. Natl. Acad. Sci. USA **93**:5517–5521.
86. **Yuan, C. X., M. Ito, J. D. Fondell, Z. Y. Fu, and R. G. Roeder.** 1998. The TRAP220 component of a thyroid hormone receptor-associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-dependent fashion. Proc. Natl. Acad. Sci. USA **95**:7939–7944.
87. **Zhang, J. J., U. Vinkemeier, W. Gu, D. Chakravarti, C. M. Horvath, and J. E. Darnell, Jr.** 1996. Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling. Proc. Natl. Acad. Sci. USA **93**:15092–15096.

Rational discovery of novel nuclear hormone receptor antagonists

Matthieu Schapira*†, Bruce M. Raaka‡, Herbert H. Samuels‡, and Ruben Abagyan*†§

*Structural Biology, Skirball Institute of Biomolecular Medicine, and ‡Division of Molecular Endocrinology, Departments of Medicine and Pharmacology, New York University School of Medicine, 550 First Avenue, New York, NY 10016

Edited by Peter G. Schultz, The Scripps Research Institute, La Jolla, CA, and approved December 7, 1999 (received for review October 20, 1999)

Nuclear hormone receptors (NRs) are potential targets for therapeutic approaches to many clinical conditions, including cancer, diabetes, and neurological diseases. The crystal structure of the ligand binding domain of agonist-bound NRs enables the design of compounds with agonist activity. However, with the exception of the human estrogen receptor-α, the lack of antagonist-bound "inactive" receptor structures hinders the rational design of receptor antagonists. In this study, we present a strategy for designing such antagonists. We constructed a model of the inactive conformation of human retinoic acid receptor-α by using information derived from antagonist-bound estrogen receptor-α and applied a computer-based virtual screening algorithm to identify retinoic acid receptor antagonists. Thus, the currently available crystal structures of NRs may be used for the rational design of antagonists, which could lead to the development of novel drugs for a variety of diseases.

Members of the nuclear hormone receptor (NR) family are under the control of a wide variety of hormones and ligands, such as steroids, retinoids, thyroid hormone, 1,25-dihydroxy-vitamin D3, and prostanoids. Many of these NRs are potential targets for the therapy of a variety of diseases: antagonists of estrogen receptor-α (ERα) (e.g., tamoxifen) are clinically used for the treatment of breast cancer (1) whereas retinoic acid receptor (RAR) agonists and antagonists block the growth of a number of neoplastic cells including breast tumor cells (2, 3). Agonists for retinoid X receptors (RXRs) and peroxisome proliferator-activated receptor γ (PPARγ) are potential candidates for use in the treatment of cancer and diabetes (PPARγ is the receptor for the antidiabetic drug thiazolidinedione) (4–7), whereas Nurr1 ligands may be useful for treatment of Parkinson's disease (8). Thus, designing molecules that selectively activate or inhibit specific NRs is of considerable biological significance and will likely have the potential for use in important clinical applications.

The crystal structures of the ligand binding domain (LBD) of many members of the NR family recently have been solved, and the ligand-dependent structural changes involved in transcriptional activation have been clarified, enabling the structure-based design of specific agonists (9, 10). Recent studies on ERα also have shed light on the LBD structural changes mediated by NR antagonists (11, 12): ERα agonists and antagonists superimpose well and engage in a very similar network of hydrophobic and electrostatic contacts with the receptor. However, in the agonist-bound conformation, the C-terminal helix H12 sits like a lid on top of the ligand (11) (a similar observation was made for virtually all of the NR LBD structures solved so far; ref. 9). In contrast, the two ERα antagonists present a protruding arm that is not compatible with the "closed lid" conformation (11, 12) (Fig. 1*A*). As a result, helix H12 is pushed away from the ligand binding site and relocates in the coactivator-binding pocket of the receptor (Fig. 1*B*) (11). Moreover, the LxxML motif (where L is a leucine, M a methionine, and x any residue) of the ERα helix H12 mimics, and probably competes with, a LxxLL helical peptide found in a wide variety of coactivator proteins. The alignment of the LBD of various NRs (13) suggests that a common structural mechanism would be for the antagonists to induce the relocation of helix H12 into the hydrophobic coactivator-binding groove of the receptor. The observation that the progesterone receptor antagonist RU486 superimposes with the natural hormone progesterone, but presents a protruding arm similar to that of tamoxifen (14, 15) provides support for the universality of this mechanism of antagonistic activity.

Our goal in this study is to provide further evidence for this hypothesis by building a model of the antagonist-bound conformation of RARα, a NR that plays an important role in the differentiation and proliferation of a wide variety of cell types and for which only the agonist bound conformation is known (16–18), and to rationally and rapidly identify new antagonists for this receptor. We built a model of the antagonist-bound structure of RAR, based on the ERα/tamoxifen complex (12). The model was used for the virtual screening of a database of ≈150,000 available compounds, and antagonist candidates were tested *in vitro*. Two novel antagonists and a novel agonist were discovered. The ligands were specific for RAR, confirming the validity of our model and the potential therapeutic application of our strategy.

Materials and Methods

Building of the Model of Antagonist-Bound RAR. A helical peptide PLIREMLENP corresponding to helix H12 of RARγ was docked into the putative coactivator binding pocket of another RARγ molecule. We hypothesized that the IxxML motif contacts the coactivator binding site of the receptor, and an automatic docking procedure was carried out toward this site, with flexible protein and peptide side chains, according to a biased probability Monte Carlo energy minimization procedure (19, 20). Two critical features of the interaction between the LBDs of NRs and their coactivators were used to carry out the docking: (*i*) The "charge clamp," initially observed in the complex between SRC-1 and peroxisome proliferator-activated receptor γ (21), where a conserved glutamate (E414 in RARγ) and lysine (K246 in RARγ) at opposite ends of the hydrophobic cavity of the receptor contact the backbone of the coactivator's LxxLL box, enabled the orientation of the helical peptide. (*ii*) The finding that the leucines of the LxxLL motif of SRC-1 are buried in the hydrophobic cavity of the receptor determines which side of the helix faces the receptor. Here, the isoleucine, methionine, and leucine of the IxxML motif were buried in the binding site of RARγ. Loose distance restraints were set between the charge

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: NR, nuclear hormone receptor; RAR, retinoic acid receptor; ER, estrogen receptor; LBD, ligand binding domain; RXR, retinoid X receptor; CAT, chloramphenicol acetyltransferase.

†To whom reprint requests should be addressed. E-mail: abagyan@scripps.edu or schapira@saturn.med.nyu.edu.

§Present address: Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, MB-37, La Jolla, CA 92037.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

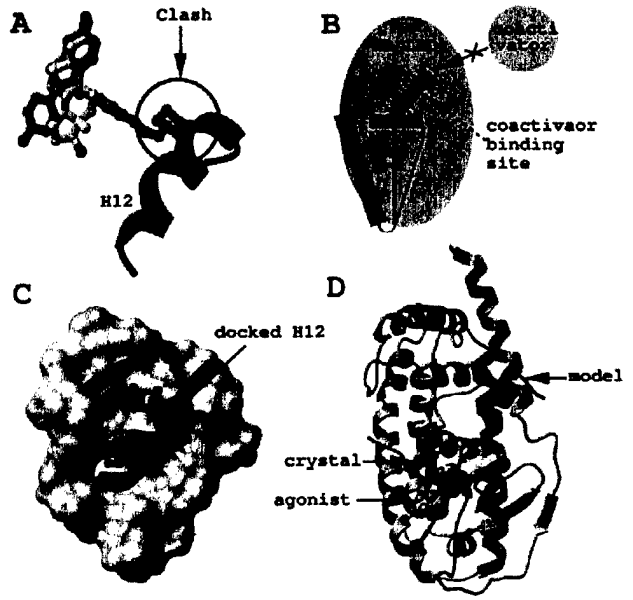


Fig. 1. Modeling of the antagonist-bound structure of RAR. Agonist (white) and antagonist (cyan) superimpose in the binding pocket of ERα, but the antagonist presents an additional protruding arm that pushes helix 12 (H12, green) away (*A*). As a result, H12 relocates in the coactivator binding pocket of the receptor (H12, red) (*B*). Based on the ERα structure, helix H12 of RARγ (red) was docked to the coactivator binding pocket of the RARγ-LBD (critical hydrophobic residues are displayed in magenta) (*C*), and the C terminus of the protein was remodeled from its agonist-bound conformation (green) to its antagonist-bound conformation (red) (*D*).

clamp of the receptor (21) (i.e., E414 and K246) and backbone nitrogen and oxygens of the peptide (nitrogen of the isoleucine on one end, and carbonyl of the methionine, leucine, and asparagine in the MLEN motifs, respectively). The energy of the complex was minimized in the internal coordinate space by using the modified ECEPP/3 potentials. The subset of the variables minimized with the ICM method (19, 20, 22, 23) included the side chains of the receptor, six positional variables of the helix, and the side-chain torsion angles of the helix.

After the ICM docking procedure, we built a model of antagonist-bound RARγ. The structure of the receptor was kept rigid but for the side chains and backbone of the 25 C-terminal residues (corresponding to the last 10 residues of helix H11, the loop from H11 to H12, and H12), and for the side chains of the putative coactivator binding site (within 6 Å of the previously docked helical peptide). Tethers then were set between the C terminus of the receptor and the corresponding residues of the docked helical peptide, and the energy of the receptor was minimized by a stochastic global energy optimization in the internal coordinate space (22, 23).

The last step was, from the resulting model of antagonist-bound RARγ, to derive the structure of the antagonist-binding pocket of RARα: the three nonidentical residues in the vicinity of the binding pocket (A234, M272, and A397) were changed to the RARα isoform (S234, I272, and V397, respectively) and energy-minimized. Another possibility would have been to introduce the mutations before remodeling the C terminus of the receptor. We preferred to proceed as described here to preserve the integrity of the receptor during the critical remodeling of the C-terminal end.

Receptor-Ligand Docking. An initial docking was carried out with a grid potential representation of the receptor and flexible ligand (24). The resulting conformation then was optimized with a full atom representation of the receptor, flexible receptor side chains, and flexible ligand, by an ICM stochastic global optimization algorithm as implemented in the MolSoft ICM 2.7 program (23, 24).

Screening of a Virtual Library of Compounds. The flexible-ligand/grid-potential-receptor docking algorithm (23, 24) was carried out automatically on the Available Chemicals Directory library of 153,000 available chemical compounds (MDL Information Systems, San Leandro, CA). The screening took less than a month on 10 194-MHz IP25 processors. Each compound was assigned a score, according to its fit with the receptor, which took into account continuum as well as discreet electrostatics, hydrophobicity, and entropy parameters (25). The distribution of the compounds according to their score is presented at http://abagyan.scripps.edu/PNAS/MS2000/. All compounds scoring better (i.e., lower) than −32 were screened further for the number of hydrogen bonds engaged with the receptor. The 134 compounds that made at least two hydrogen bonds with the receptor were preselected. The 609 compounds scoring better than −37 also were preselected, regardless of the hydrogen bonding network. This preselection pool then was further minimized with a full atom representation of the receptor, as described above. The quality of the fit of the 500 best-scoring compounds then was visually estimated, and 32 compounds were selected for biological testing. These compounds are not necessarily the ones with the best final scores, but the ones we thought, after careful visual inspection, presented the best characteristics, such as Van der Waals fit or hydrogen bonding (see http://abagyan.scripps.edu/PNAS/MS2000/).

It occurred to us that during the selection by the MolSoft virtual screening procedure, it was preferable to set up an initial cut-off value poorly selective (i.e., −32) to recover a large pool of preselected compounds and to apply to this pool subsequent screens specific for the system, such as number of hydrogen bonds (used here) or presence of a hydrogen bond acceptor (for example) at a specific point of space. As a result, we derived the value −32 as a good initial threshold (this value generates an initial pool of 3,000–4,000 compounds).

Biological Activity of the Antagonist and Agonist Candidates. HeLa cells were transfected by calcium phosphate precipitation using 1 μg of the Gal4-responsive chloramphenicol acetyltransferase (CAT) reporter pMC110 and 1 μg of Gal4-hRARα-LBD or 1 μg of Gal4-hRXRβ-LBD. Studies also were performed with the three wild-type hRAR isoforms (hRARα, hRARβ, and hRARγ) by using a ΔMTV-IR-CAT reporter as described (26, 27). Cell cultures were supplemented with indicated ligands immediately after addition of the calcium phosphate/DNA precipitate. Media and ligands were replaced after 24 h, and cells were harvested and essayed for CAT activity 24 h later.

Results

Modeling of the RAR Antagonist Binding Pocket. The x-ray structure of RARγ bound to the agonist all-trans RA is available (18); however, the conformation of the receptor bound to an antagonist is not known. We used the observations made from the structure of ERα bound to an agonist, 17β-estradiol (11), and two antagonists, tamoxifen and raloxifene (11, 12), to build a model of antagonist-bound RAR (Fig. 1 *A* and *B*). We docked helix H12 of RARγ into the putative coactivator binding pocket of the receptor as described (27) (see *Materials and Methods* for details) (Fig. 1*C*) and remodeled the 25 C-terminal residues, starting near the end of helix 11, through an extensive global energy minimization procedure (Fig. 1*D*).

Docking of Known RAR Antagonists into the Modeled Receptor. A few RAR antagonists have been described in the literature; and

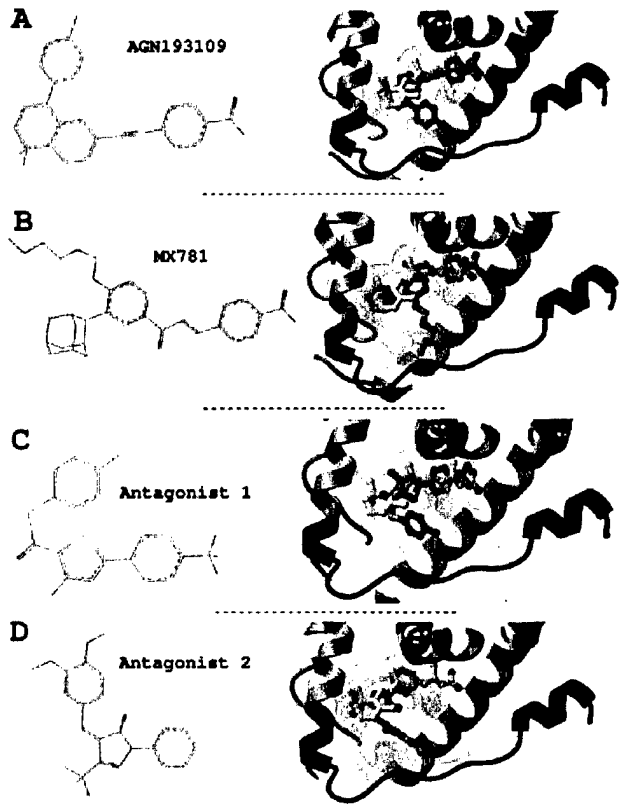


Fig. 2. RAR antagonists. Two known antagonists (*A* and *B*) and two novel antagonists (*C* and *D*). (*Left*) Chemical structure. (*Right*) Conformation docked into the receptor (part of the receptor is displayed as a ribbon representation, and the binding pocket boundary is displayed in yellow). Cyan, carbons; red, oxygen; blue, nitrogen; magenta, fluorine; yellow, sulfur. Hydrogens are not represented for clarity.

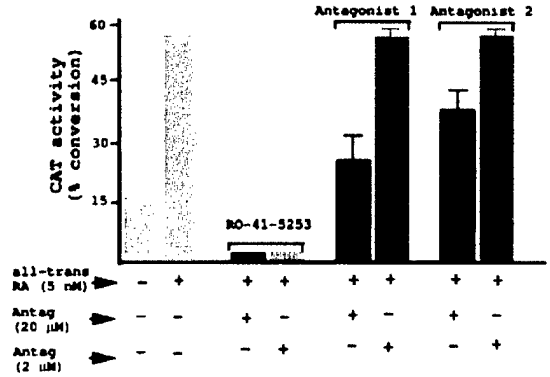


Fig. 3. Functional assays of the novel antagonists. HeLa cells were transfected with a Gal4-hRARα-LBD expression vector and a Gal4-CAT reporter gene (results were similar in studies using the three hRAR isoforms). The cells were incubated with 5 nM all-trans RA to stimulate CAT activity, and the effect of each antagonist on inhibiting CAT was examined at 2 and 20 μM concentration (the known antagonist RO-41–5253 was used as a positive control).

several of them are serious candidates for cancer therapy (2, 28). A well-characterized ligand is AGN193109, which inhibits the three RAR isoforms at nanomolar concentrations (29). Another very potent antagonist is MX781, which is effective against ERα-positive and -negative breast cancer cells, with no apparent toxicity (2). The activity of these two ligands has been presented in detail, but no structural information has been reported on their mode of interaction with the receptor. We built a model of RARγ complexed either with AGN193109 or MX781, by using our flexible docking algorithm (24) (Fig. 2 *A* and *B*). In both cases, the antagonist superimposed with the agonist all-trans RA. As observed for ERα, the antagonists also presented a protruding arm, which was absent in RAR agonists. Very importantly, this protruding arm coincided exactly with the single opening in the ligand binding pocket of our modeled receptor, generated by the displacement of helix H12 (Fig. 2 *A* and *B*), and made stabilizing hydrophobic contacts with the protein. It is very unlikely that this perfect fit, observed for both antagonists, was fortuitous. On the contrary, this feature mimics the inactivation mechanism revealed by the crystal structure of ERα bound to tamoxifen and raloxifene. Therefore, our docking results of AGN193109 and MX781 very strongly suggest that: (*i*) the structural mechanisms of antagonist activity for ERα are shared by other NRs, and (*ii*) our model of the RAR antagonist binding pocket could be used to design novel antagonists.

Screening of a Virtual Library and Discovery of Novel RAR Antagonists. High throughput functional screening currently is the most used method for the discovery of receptor-specific ligands. Although efficient, it requires the physical availability and management of hundreds of thousands of chemical compounds. In the present work, we used a virtual library composed of the predicted structure of more than 150,000 available compounds (see *Materials and Methods*). Each compound was automatically docked in a grid representation of the modeled RARα antagonist binding pocket. Five grid potentials carried information on the shape, hydrophobicity, electrostatics, and hydrogen-bonding availability of the receptor, and enabled a rapid docking simulation (24, 25). RARα was selected over the other two isoforms (RARβ and RARγ) because recent data suggests it could be a medically more relevant target (28). After an automatic selection procedure with flexible ligands, and optimization of the selected candidates with flexible protein side chains (see *Materials and Methods* for details), 32 compounds were considered as potential antagonists of RARα and ordered.

To test these compounds *in vitro*, HeLa cells were transfected with a Gal4-hRARα-LBD expression vector and a Gal4-CAT reporter gene (26). Studies also were performed with the three wild-type hRAR isoforms and a ΔMTV-IR-CAT reporter (26, 27). These gave similar results as those found with Gal4-hRARα-LBD (data not shown). The cells were incubated with all-trans RA to stimulate CAT activity, and the effect of each antagonist candidate on inhibiting CAT stimulation by all-trans RA was examined. Possible toxicity of the compounds was deduced from the amount of cellular protein extract after 2 days of incubation. Two antagonist candidates inhibited CAT activity by 55% and 33% at 20 μM with no apparent toxicity (Fig. 3). The Gal4-hRARα activity illustrated in Fig. 3 was equivalent for the other two RAR isoforms (data not shown). No inhibition was observed when CAT expression was under the control of a Gal4-mRXRβ-LBD fusion construct, indicating that: (*i*) the antagonists are specific for RAR, and (*ii*) the inhibition is caused by an interaction with the Gal4-RAR-LBD fusion protein and does not result from some nonspecific effect on CAT activity (data not shown).

The two RAR antagonists dock into the ligand binding pocket of the receptor (Figs. 2 *C* and *D* and 4). As observed for AGN193109 and MX781, they fit in the same binding pocket as the natural agonist all-trans RA, but present an additional arm, which protrudes out of the pocket. Antagonist 1 has a tri-fluoro group where the retinoid receptor ligands usually carry a carboxylate group (in antagonist 2, the corresponding domain is truncated). In our model, antagonist 2 engages in a hydrogen bond with Ser-234 of the hRARα (Fig. 4*B*). However, the S234A

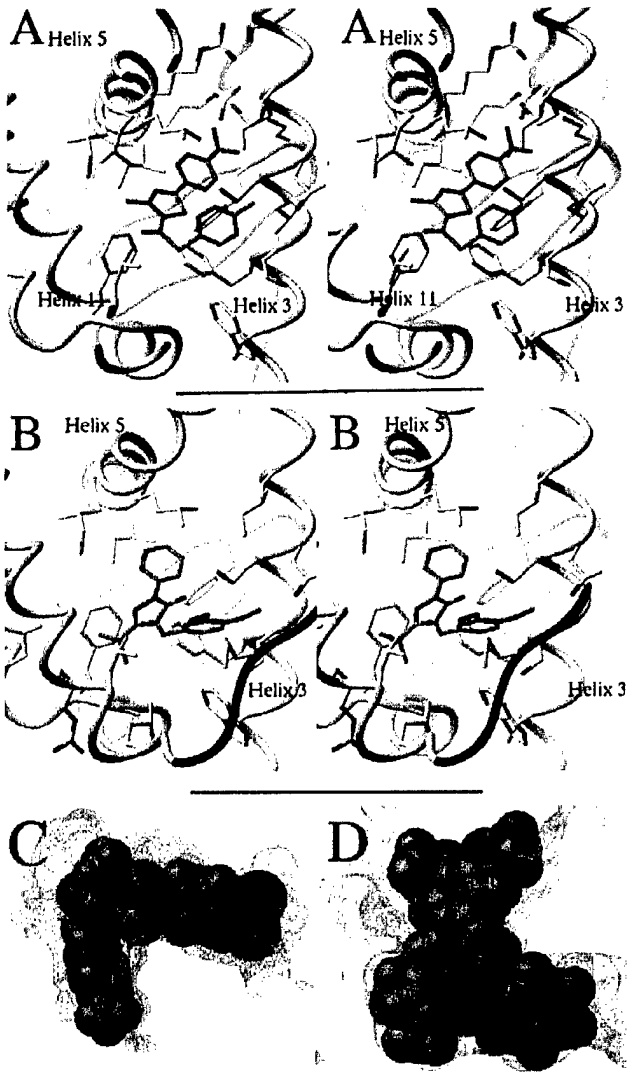


Fig. 4. Novel RAR antagonists. (*A* and *B*) Stereo representation of antagonists 1 and 2 docked into the binding site of the receptor. The ligands make extensive hydrophobic interactions with residues from helix 3, helix 5, and helix 11. Antagonist 2 (*B*) is engaged in an additional hydrogen bond with Ser-234 of helix 3 and contacts the remodeled C terminus (red) at Pro-405. (*C* and *D*) The fit of antagonists 1 and 2 into the receptor binding pocket is shown.

mutation in the other two isoforms does not alter the ligand antagonist activity, suggesting that this hydrogen bond is not essential for the interaction. An obvious way to increase the affinity of these antagonists would be to substitute the tri-fluoro group by a carboxylate in antagonist 1 or elongate and add a carboxylate to antagonist 2, which would result in more stabilizing interactions with two conserved arginines of the receptor. However, the purpose of this work is to provide evidence that the rational design of antagonists from the model of the inactive receptor is feasible and not to optimize the affinity of the compounds. The *in vitro* functional assays provide evidence that our modeling scheme is relevant and can be used to design novel antagonists of NRs.

We applied the same strategy to discover agonists, by using the crystal structure of the active conformation of RARγ (18), and could discover three novel agonists 10–25% active at 200 nM and fully active at 20 μM, of 30 compounds tested (data not shown).

Screening of a Database of Known Ligands. To assess the quality of our setup of the ICM screening algorithm (23), we built a small virtual database made up of antagonists and agonists for different members of the NR family (Table 1). We screened this database with our model of antagonist-bound RAR, as we did for the Available Chemicals Directory database. The screening was repeated four times, to test the reproducibility of our method. Table 1 shows that for each ligand the score varies a lot from one screening to the other. This finding reflects the generation of different ligand conformations from one docking simulation to another (data not shown) and represents the limitation of our method, as discussed below.

Table 1 lists as "selected" the ligands that met with the criteria for preselection and final inspection during the Available Chemicals Directory screening (i.e., score better than -37 or score better than -32 and at least two hydrogen bonds with the receptor; see *Materials and Methods* for details). Seven of the nine known RAR ligands (i.e., $\approx$80%) and one of the six non-RAR ligands (i.e., $\approx$16%) were selected. The fact that RAR agonists, as well as antagonists, produced good scores was expected, because the binding pocket used for the screening is equivalent to the agonist binding pocket, with an additional opening generated by the remodeling of the C terminus of the receptor. The two false negatives, AGN193836 and Ro415253, were missed because of steric clashes, as discussed below. Antagonist 1 was not found either, reflecting its rather low affinity for the receptor. It is important to underline here that we do not expect to detect all of the true binders. The algorithm was rather designed to minimize the number of false positives, which correlates with the number of unnecessary *in vitro* experiments (25).

In that respect, the presence of one false positive of six nonbinders could be alarming, because such a ratio would represent about 25,000 false positives of a database of 150,000 compounds. However, the binding pockets of the NRs represented in this database are close in size and shape; as a result, the database used for this benchmark was composed of molecules presenting strong similarities with RAR ligands. Therefore, we believe this ratio is not representative. The fact that we needed to test only 32 molecules to discover three novel RAR ligands confirms this assumption.

Next, we tried to address why some ligands, such as Ro415253, were repeatedly missed by our screening algorithm (Ro415253 was still not selected after 10 docking simulations, data not shown). We hypothesized that the ligand could not fit into the potential maps generated from our model and carried out a docking simulation with a full atom representation of the receptor, according to a Monte Carlo energy minimization of the complex, with both flexible ligand and flexible receptor side chains (24). This docking simulation produced a solution were the ligand fits into the binding pocket; the core of the ligand (from the carboxylate to the internal sulfone) superimposes with agonists such as all-trans RA, whereas the alkyl arm sticks out of the pocket, as previously described for the other antagonists (data not shown). The conformation of several receptor side chains was modified during the docking simulation, to accommodate the size of the ligand, and this solution would not have been found with rigid side chains. This finding suggests that Ro415253 could not fit into the potential maps generated from the original receptor conformation, which we used for the screening. We generated a new series of potential maps from the optimized receptor structure and screened the small database of known ligands with these maps four times as above (Table 1). The score assigned to Ro415253 was twice lower (i.e. better) than the threshold. Surprisingly, this new series of potential maps totally eliminated the presence of both false positive and false negative (all RAR ligands and only RAR ligands were selected).

Table 1. Control screening of known NR ligands

Ligand	Activity	Score 1	Score 2	Score 3	Score 4	Selected	Binding	References
First series								
AGN193836	RAR_agonist	−19.9	−9.04	−20.6	−19.7	−	+	(33)
ATRA	RAR pan-agonist	−46.4	−41	−41.7	−41.	+	+	(34)
Ro415253	RAR_antagonist	−25.5	−22.	−28.3	−28.6	−	+	(28)
MX781	RAR antagonist	−28.	−23.9	−27.1	−36.4	+	+	(2)
CD2366	RAR pan-antagonist	−28.5	−23.3	−30.9	−32.3	+	+	(34)
Targretin	RXR pan-agonist	−17.9	−18.1	−19.1	−18.6	−	−	(4)
SR11203	RXR pan-agonist	−27.5	−27.	−27.	−27.2	−	−	(34)
Tamoxifen	ER modulator	−29.3	−27.5	−29.8	−28.3	−	−	(23)
Raloxifene	ER modulator	−23.4	−20.8	−26.7	−34.6	+	−	(22)
RU486	Progest Rec antag.	−21.2	−21.3	−21.4	−21.3	−	−	(25)
9cisRA	RAR/RXR agonist	−32.5	−32.6	−32.9	−16.9	+	+	(34)
AGN193109	RAR pan-antagonist	−39.2	−56.	−57.4	−39.4	+	+	(29)
AGNpartia	RAR partial agonist	−54.4	−54.3	−49.5	−29.1	+	+	(29)
Am580	RAR_agonist	−34.2	−34.4	−34.8	−34.5	+	+	(34)
EM652	ER antagonist	−27.	−27.4	−21.7	−28.8	−	−	(35)
Antagonist 1	Novel RAR antag.	−28.5	−28.1	−28.7	−28.8	−	+	(35)
Antagonist 2	Novel RAR antag.	−27.6	−38.9	−40.2	−26.3	+	+	(35)
Second series								
AGN193836	RAR_agonist	−37.2	−36.5	−36.7	−35.3	+	+	(33)
ATRA	RAR pan-agonist	−51.7	−52.6	−51.8	−52.0	+	+	(34)
Ro415253	RAR_antagonist	−28.9	−24.4	−39.0	−46.6	+	+	(28)
MX781	RAR antagonist	−45.3	−48.0	−40.2	−45.6	+	+	(2)
CD2366	RAR pan-antagonist	−50.7	−50.8	−29.3	−29.3	+	+	(34)
Targretin	RXR pan-agonist	−25.4	−23.0	−22.2	−31.0	−	−	(4)
SR11203	RXR pan-agonist	−28.2	−22.7	−22.1	−27.5	−	−	(34)
Tamoxifen	ER modulator	−26.4	−24.6	−30.3	−23.4	−	−	(23)
Raloxifene	ER modulator	−15.6	−23.7	−18.4	−17.4	−	−	(22)
RU486	Progest Rec antag.	−21.4	−20.6	−20.3	−20.1	−	−	(25)
9cisRA	RAR/RXR agonist	−38.8	−39.5	−33.5	−38.7	+	+	(34)
AGN193109	RAR pan-antagonist	−55.1	−55.5	−41.2	−54.8	+	+	(29)
AGNpartia	RAR partial agonist	−61.4	−61.3	−61.4	−61.0	+	+	(29)
Am580	RAR_agonist	−46.6	−47.2	−46.6	−46.5	+	+	(34)
EM652	ER antagonist	−26.3	−23.1	−23.7	−27.3	−	−	(35)
Antagonist 1	Novel RAR antag.	−32.1	−32.1	−31.7	−31.6	+	+	(35)
Antagonist 2	Novel RAR antag.	−33.3	−29.7	−33.8	−33.8	+	+	(35)

First series: A similar screening as the one performed on the ACD database was carried out four times on a small database made of known RAR antagonists, agonists, as well as ligands for other NRs and the two novel RAR antagonists. The ligands that met at least once with the criteria for selection used during the ACD screening are listed as Selected. The ligands that are experimentally binding to RAR are listed as Binding. Second series: Screening of known ligands after adjustment of the receptor's binding pocket conformation. The RAR antagonist Ro415253 was docked into our model of antagonist-bound RAR with flexible receptor side chains and ligand. The resulting receptor conformation was used for a novel screening.

Discussion

In this study, we presented a strategy for the discovery of antagonists, as well as agonists, for NRs, which are very important targets for drug design. An important aspect of our approach was to exclude any preconceived pharmacophore bias from our database screening. Most drug design strategies impose chemical constraints on the selected molecule to conserve the functional groups believed to be most important in existing ligands, preventing the discovery of novel ligand types. In the present work, we avoided pharmacophore constraints thanks to a robust flexible docking program and scoring function: the only filters used for screening were a good fit with the receptor and reasonable bioavailability parameters (30). As a result, we discovered novel original ligands that could be further optimized into potent RAR-selective antagonists and agonists.

A limitation of our method, which leaves room for further improvement, is that a compromise must be made between the time allocated for each ligand (less than 2 min on one processor here) and the reliability of the sampling of the conformational space. Indeed, Table 1 shows that four runs for each ligand are necessary to minimize efficiently missed hits (the remaining missed positives were not selected because of inappropriate receptor side-chain conformations and not because of an insufficient sampling). Improvement of the computing power, the docking algorithm, and the scoring function all could result in a more robust virtual database screening.

Another drawback is that the conformation of the receptor is not necessarily unique, but can vary from one ligand to another. As a result, a ligand that fits in receptor conformation A will never be found if receptor conformation B is used for the screening. The case of Ro415253 illustrates this issue well: this known antagonist was never selected, even after 10 trials, because the binding pocket used for the screening was too narrow. The potential maps used for the screening have a smoother van der Waals profile than the atomic representation of the receptor; as a result, the maps are more tolerant regarding steric clashes with the ligand. However, the degree of tolerance is limited and cannot accommodate important conformational changes of the receptor side chains (or backbone, obviously). When new potential maps generated from a model of RAR bound to Ro415253 were used for screening, the three RAR ligands missing from the first screening were selected (Table 1). This finding confirms that the initial conformation of the

receptor prevented the selection of, or reduced the chances of selecting, some known RAR ligands. The false positive raloxifene (Table 1) was making extensive van der Waals interactions with the narrow RAR binding pocket, which compensated for the lack of stabilizing electrostatic interactions. However, in the new conformation of the receptor (Table 1), the binding pocket is wider and the fit not as tight. As a result, raloxifene was not selected. This observation emphasizes, if necessary, that virtual screening is very sensitive to the conformation of the receptor.

In that respect, it is interesting to note that the topology of the remodeled C-terminal loop is probably not unique, and that the conformation used to generate the receptor potential maps was one among many others. It is therefore legitimate to wonder whether novel antagonists could not be discovered as efficiently from a structure of the receptor where the C terminus, instead of being remodeled, was truncated. This brings up a fundamental question: is the role of antagonists only to antagonize the "closed lid" conformation where helix H12 sits on top of the ligand binding pocket, or are they also stabilizing the inactive conformation of the receptor? It is important to keep it mind that the C-terminal tail of RAR (as well as for other NRs) is a very dynamic entity when no ligand is bound to the receptor and probably oscillates between active and inactive conformations. Once bound in the ligand binding pocket, agonists contact the H12 helix and lock the receptor in its coactivator-binding conformation. Likewise, it is reasonable to speculate that antagonists would contact the C-terminal tail of the receptor and stabilize the inactive state. However, it is probable that the conformation of the receptor varies from one ligand to another; indeed, recent results on ERα show that different ligands induce distinct conformational change of the receptor (31). We used the crystal structure of ERα bound to tamoxifen to build our model of inactive RAR and could find two specific antagonists, one of which contacts the remodeled tail of the receptor. Although the conformation we used for the C-terminal tail was probably not the only possible one, we believe that its presence was important to bias the screening toward compounds that actually do contact the flexible arm of RAR, as well as to impose a reasonable boundary on the antagonist binding pocket, and prevent the ligands from drifting out of the pocket during the docking simulations.

An important point was to demonstrate that we could discover novel antagonists for a NR other than ERα, provided that the structure of the agonist-bound active form of the protein was known. Rational design of ligands from a model of a receptor is thought by many to yield very low success rates. The present study demonstrates that this strategy can be successfully undertaken with appropriate biological systems and robust modeling tools. Moreover, targeting models of diverse members of the NR family could be further justified by the wealth of structural and sequence information (9, 13), as well as the finding that NR family members share similar mechanisms of transcriptional activation and inhibition (9).

The recent publication of the crystal structures of medically relevant receptor targets, such as peroxisome proliferator-activated receptor γ (21), RAR (18), RXR (32), ERα (11), or progesterone receptor (15), has created an exciting opportunity for the discovery of novel ligands. This study demonstrates that the rational design of both antagonists and agonists, by using computer-generated models based on these structures, is possible.

We thank M. Totrov for helpful discussion. We thank MolSoft LLC for making the latest version of the ICM program available for this research project. This research was supported by Department of Defense Grant DAMD179818133, a Kaplan Comprehensive Cancer Center grant from New York University Medical Center, National Institutes of Health Grant GM5541801, and Department of Energy Grant DEFG0296ER62268 to M.S. and R.A., and National Institutes of Health Grant DK16636 and New York State Empire Award C015710 to H.H.S.

1. Dees, E. C. & Kennedy, M. J. (1998) *Curr. Opin. Oncol.* **10**, 517–522.
2. Fanjul, A. N., Piedrafita, F. J., Al-Shamma, H. & Pfahl, M. (1998) *Cancer Res.* **58**, 4607–4610.
3. Shiohara, M., Dawson, M. I., Hobbs, P. D., Sawai, N., Higuchi, T., Koike, K., Komiyama, A. & Koeffler, H. P. (1999) *Blood* **93**, 2057–2066.
4. Bischoff, E. D., Moon, T. E., Heyman, R. A. & Lamph, W. W. (1998) *Cancer Res.* **58**, 479–484.
5. Mukherjee, R., Davies, P. J., Paterniti, J. R., Jr. & Heyman, R. A. (1997) *Nature (London)* **386**, 407–410.
6. Spiegelman, B. M. (1998) *Diabetes* **47**, 507–514.
7. Elstner, E., Muller, C., Koshizuka, K., Williamson, E. A., Park, D., Asou, H., Shintaku, P., Said, J. W., Heber, D. & Koeffler, H. P. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 8806–8811.
8. Zetterstrom, R. H., Solomin, L., Jansson, L., Hoffer, B. J., Olson, L. & Perlmann, T. (1997) *Science* **276**, 248–250.
9. Moras, D. & Gronemeyer, H. (1998) *Curr. Opin. Cell Biol.* **10**, 384–391.
10. Klaholz, B. P., Renaud, J. P., Mitschler, A., Zusi, C., Chambon, P., Gronemeyer, H. & Moras, D. (1998) *Nat. Struct. Biol.* **5**, 199–202.
11. Brzozowski, A. M., Pike, A. C., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G. L., Gustafsson, J. A. & Carlquist, M. (1997) *Nature (London)* **389**, 753–758.
12. Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A. & Greene, G. L. (1998) *Cell* **95**, 927–937.
13. Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D. & Gronemeyer, H. (1996) *Nat. Struct. Biol.* **3**, 87–94.
14. Cadepond, F., Ulmann, A. & Baulieu, E. E. (1997) *Annu. Rev. Med.* **48**, 129–156.
15. Williams, S. P. & Sigler, P. B. (1998) *Nature (London)* **393**, 392–396.
16. Fitzgerald, P., Teng, M., Chandraratna, R. A., Heyman, R. A. & Allegretto, E. A. (1997) *Cancer Res.* **57**, 2642–2650.
17. Giannini, G., Dawson, M. I., Zhang, X. & Thiele, C. J. (1997) *J. Biol. Chem.* **272**, 26693–26701.
18. Renaud, J. P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H. & Moras, D. (1995) *Nature (London)* **378**, 681–689.
19. Totrov, M. & Abagyan, R. (1994) *Nat. Struct. Biol.* **1**, 259–263.
20. Strynadka, N. C., Eisenstein, M., Katchalski-Katzir, E., Shoichet, B. K., Kuntz, I. D., Abagyan, R., Totrov, M., Janin, J., Cherfils, J., Zimmerman, F., *et al.* (1996) *Nat. Struct. Biol.* **3**, 233–239.
21. Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K. & Milburn, M. V. (1998) *Nature (London)* **395**, 137–143.
22. Abagyan, R. & Totrov, M. (1994) *J. Mol. Biol.* **235**, 983–1002.
23. MolSoft (1998) ICM 2.7 *Program Manual* (MolSoft, San Diego).
24. Totrov, M. & Abagyan, R. (1997) *Proteins* Suppl. **1**, 215–220.
25. Totrov, M. & Abagyan, R. (1999) in *Proceedings of the Third Annual International Conference on Computational Molecular Biology*, April 1999, Lyon, France (ACM Press, New York), pp. 37–38.
26. Qi, J. S., Desai-Yajnik, V., Greene, M. E., Raaka, B. M. & Samuels, H. H. (1995) *Mol. Cell. Biol.* **15**, 1817–1825.
27. Li, D., Desai-Yajnik, V., Lo, E., Schapira, M., Abagyan, R. & Samuels, H. H. (1999) *Mol. Cell. Biol.* **19**, 7191–7202.
28. Toma, S., Isnardi, L., Raffo, P., Riccardi, L., Dastoli, G., Apfel, C., LeMotte, P. & Bollag, W. (1998) *Int. J. Cancer* **78**, 86–94.
29. Chandraratna, R. A. (1998) *J. Am. Acad. Dermatol.* **39**, S149–S152.
30. Lipinski, C. A., Lombard, F., Dominy, B. W. & Feeney, P. J. (1997) *Adv. Drug Delivery Rev.* **23**, 3–25.
31. Paige, L. A., Christensen, D. J., Gron, H., Norris, J. D., Gottlin, E. B., Padilla, K. M., Chang, C. Y., Ballas, L. M., Hamilton, P. T., McDonnell, D. P. & Fowlkes, D. M. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 3999–4004.
32. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. (1995) *Nature (London)* **375**, 377–382.
33. Teng, M., Duong, T. T., Klein, E. S., Pino, M. E. & Chandraratna, R. A. (1996) *J. Med. Chem.* **39**, 3035–3038.
34. Sun, S. Y., Yue, P., Dawson, M. I., Shroot, B., Michel, S., Lamph, W. W., Heyman, R. A., Teng, M., Chandraratna, R. A., Shudo, K., *et al.* (1997) *Cancer Res.* **57**, 4931–4939.
35. Tremblay, A., Tremblay, G. B., Labrie, C., Labrie, F. & Giguere, V. (1998) *Endocrinology* **139**, 111–118.

In Silico Discovery of novel RAR Agonist Structures

Matthieu Schapira [1,*], Bruce M. Raaka[2], Herbert H. Samuels[2], and Ruben Abagyan[1,3,*]

1 Structural Biology, Skirball Institute of Biomolecular Medicine

2 Division of Molecular Endocrinology, Department of Medicine and Department of Pharmacology

New York University School of Medicine

550 First Avenue

New York, NY 10016, USA

3 Present Address: Department of Molecular Biology

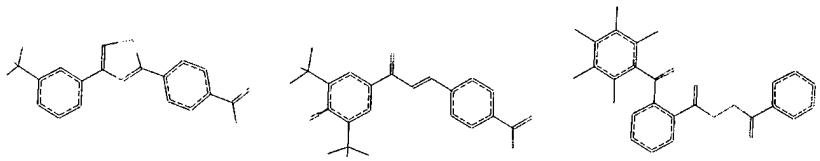
The Scripps Research Institute

10550 North Torrey Pines Road, MB-37

La Jolla, CA 92037, USA

*Correspondence should be addressed to R.A. (tel 1 619 784-8595, fax 1 619 784-8299 abagyan@scripps.edu) or M.S. (tel. 1 212 263 7817, fax 1 212 263 8951 schapira@saturn.med.nyu.edu)

Table of Contents Graphic



black: carbon, red: oxygen, blue: nitrogen, yellow: sulfur, magenta: fluoride

Abstract

Several RAR agonists have therapeutic activity against a variety of cancer types; however, unacceptable toxicity profiles have hindered the development of drugs. RAR agonists presenting novel structural and chemical features could therefore open new avenues for the discovery of leads against breast, lung and prostate cancer or leukemia. We have applied a high throughput flexible internal coordinate docking approach (ICM) to discover such compounds. A library of over 150.000 molecules was docked *in silico* to the structure of the receptor and 30 ligand candidates were tested *in vitro*. Three novel agonists were identified, one of them, agonist 3, presenting two original features: (i) a penta-methylated benzene linked by a ketone to a second aromatic ring, forming a very large hydrophobic head. (ii) a pyridine replacing a carboxylate or ester present in all RAR ligands described so far. These novel structural features may translate into improved toxicity profiles and result in the development of new ligands for cancer therapy.

Introduction

The retinoic acid receptors (RAR-α, -β, and -γ) are transcription factors regulating a variety of endocrine metabolic pathways. Unlike anti-estrogens, such as tamoxifen or raloxifene, ligands targeted against the RAR isoforms can present anticancer activity against both estrogen receptor positive and negative breast tumor cells[1]. As a result, such molecules could constitute a novel generation of drugs against breast cancer. For reasons not yet clear, both agonists and antagonists of RAR can present anti-tumor activity against breast, prostate, lung cancer or leukemia[1-7]. The development of both types of ligands could therefore have important biomedical implications. We have recently demonstrated that antagonists could be discovered rationally, based on a model of the antagonist-bound conformation of the receptor[8]. Our goal here is to discover innovative molecular structures with RAR agonist activity.

Several retinoid and non-retinoid ligands have been described, which activate one or a combination of RAR isoforms. Some of them, such as the natural ligand all-trans retinoic acid (all-trans RA), have been tested clinically, and display unacceptable side effects, such as skin dryness, cheilitis, hypertriglycemia and conjunctivitis[9-10]. However, the compounds tested so far belong to limited series of related structures. An increasing amount of data suggest that the RARα isoform, which controls the expression of RARβ, is the most relevant target for anticancer therapy[11-15]. Innovative molecules with RARα agonist activity could therefore present more favorable toxicity profile.

We applied a flexible virtual screening algorithm (Molsoft ICM, virtual screening module[16]) which rapidly docks hundreds of thousands of flexible compound structures into the ligand binding pocket of RARα and discovered three novel RAR pan-agonists. At least one ligand has unexpected structural and chemical characteristics, which could be used in the development of novel compounds for cancer prevention and therapy.

Results and discussion

We first built a model of the RARα agonist binding pocket from the crystal structure of the

RARγ ligand binding domain (RARγ LBD) / all-trans RA complex[17]. All but three amino acids in the vicinity of the ligand are conserved between the two isoforms. These three non-identical residues -A234, M272, and A397- were changed to the RARα isoform -S234, I272 and V397- and energy minimized (see "Experimental Section").

In order to address the accuracy of our model of the RARα binding pocket, we docked AM580, an RARα specific agonist[18], into the receptor. A rapid docking procedure with flexible ligand and a grid representation of the receptor was followed by an extensive Monte Carlo energy minimization with both ligand and receptor side chains flexible (see "Experimental Section" for details). The ligand superimposed well with the natural ligand all-trans RA (Figure 1a). Interestingly, Am580 does not seem to fit in the receptor binding site: the ketone oxygen of the ligand sticks out of the binding pocket, due to too close proximity of residue 234 (Figure 1b). However, in the complex with RARα, this ketone oxygen actually shares an hydrogen atom with the hydroxyl group of serine 234, and forms a stabilizing hydrogen bond, while a steric clash occurs in the other two RAR isoforms, where residue 234 is an alanine. Consequently, our model provides a rational for Am580 isoform specificity, suggesting that it is relevant and could be used as a template for the discovery of novel RARα agonist structures.

A high throughput virtual screening was carried out on the Available Chemicals Directory (MDL Information Systems, San Leandro, CA), a compound structure database of over 150,000 molecules. Each compound was automatically docked into a grid representation of RARα, as previously described for Am580, and assigned a score according to the quality of the fit[19,16]. The 5364 ligand candidates which scored better (i.e. lower) than -32 kcal/mol were pre-selected for a more refined energy minimization procedure, with flexible receptor side chains (see "Experimental Section" for details). After careful visual examination, 30 molecules were selected and purchased to be experimentally tested *in vitro*.

HeLa cells were transfected with a Gal4-hRARα-LBD expression vector and a Gal4-CAT reporter gene[20]. The cells were incubated with each ligand at 0.1, 1 and 10 μM to stimulate CAT activity. Possible toxicity of the compounds was deduced from the amount of cellular protein extract after 2 days of incubation. The percentage of conversion induced by all-trans RA at 100

nM was used as a positive control for maximum induction. Two agonist candidates (agonist 1 and agonist 2) induced 92% and 98% of the maximum CAT activity at 10 μM, and 23% and 88% of maximum CAT activity at 1 μM respectively, with no apparent toxicity (Figure 2). Agonist 2, in particular, was a rather good agonist: it still could induce 44% of maximum CAT activity at 200 nM. The Gal4-hRARα-LBD activity was similar to the other two RAR isoforms, and studies performed with the three wild type hRAR isoforms and a ΔMTV-IR-CAT[20,21] reporter confirmed the results (data not shown). The compounds could also activate a Gal4-mRXRβ-LBD fusion construct at 10 μM: 60% and 20% of maximal CAT activity was induced by Agonist 1 and 2 respectively (data not shown).

Agonist 1 and 2 both have a carboxylate group which superimposes with the carboxylate of all-trans RA, and make stabilizing hydrogen bonds with Arg 274 and 278, and the backbone nitrogen of Ser 289 (Figure 3). An additional electrostatic interaction is probably present between Ser 234 and the thiazole nitrogen of agonist 1. However, it is not necessary for the activity of the compound since residue 234 is an alanine in RARβ and RARγ, and no difference of activity is observed between the three isoforms. All other interactions between the receptor and RARα are hydrophobic. As a result, the size of the ligand and the shape complementarity with the receptor are critical for affinity and specificity. In that respect, it is interesting to note that, while agonist 1 is close in size to all-trans RA, agonist 2 is substantially larger, thanks to its bulky head composed of two tertio-butyl groups. This feature probably underlies the observation that agonist 2 is more active than agonist 1 (Figure 2).

We recently described a similar virtual screening procedure carried out on a model of antagonist-bound RAR, where the C-terminal helix H12 was translocated onto the coactivator binding site of the receptor, in an effort to discover novel RAR antagonists[8]. One of the antagonist candidates could fit into our model of active RARα, and presented substantial agonist activity: it induces 28% of maximum CAT activity at 1 μM and 73% at 10 μM (Figure 2, agonist 3). This compound is much larger than all-trans RA and agonists 1 and 2 (Figure 3). Agonist 3 did not activate RXR at concentrations tested (up to 20 μM) (data not shown).

To our knowledge, virtually all RAR agonists described in the literature have three characteristic

structural features: (i) A bulky hydrophobic head, such as the poly-methylated ring of all-trans RA or the tetra-methyl tetra-hydro naphthalene of AM580[18]. (ii) An elongated, mostly hydrophobic linker about 10 A long (10.8 A for all-trans RA). (iii) A carboxylate or ester end, which makes specific electrostatic contacts with Arg 274 and 278, and Ser 289 (Figure 3A and 3B). Figure 3 shows that both agonist 1 and agonist 2 present these characteristics. Agonist 2 is actually very close to Ch55, an already known RAR agonist[22]. However, it is worth mentioning that the tri-fluoro group at the head, and the thiazole ring in the core of agonist 1 constitute original chemical structures, for RAR ligands.

Unlike agonist 1 and agonist 2, the hydrophobic head of agonist 3 is surprisingly composed of two aromatic moieties, linked by a ketone group (Figure 3C). A bipartite head was already observed in compounds with an adamantane moiety, such as the RARγ selective agonist CD437[23]. However, the penta-methylated ring in agonist 3 and the ketone linking the two aromatic moieties constitute to our knowledge an entirely novel structural feature, which fits tightly into the RAR binding pocket, and makes extensive Van der Waals and hydrophobic contacts with the receptor. Another unexpected feature of agonist 3 is that its linker region is hydrophilic which is unfavorable to the stability of the complex (most of the polar groups of the linker are surrounded by hydrophobic residues), and decreases the affinity of the ligand for the receptor. Consequently, some chemical modifications of agonist 3, such as replacing the nitrogens of the linker by carbons, could result in a molecule with improved affinity for the receptor.

The third and most surprising characteristic of agonist 3 is that, to our knowledge, it is the first RAR ligand described so far which lacks a carboxylate group (some heteroarotinoids have been described with a methyl or ethyl ester at the carboxylate position, but this group might be processed *in vivo* into the carboxylate[24,25]. Moreover, docking experiments suggest that the nature of the interaction with arginine 278 of the receptor is conserved (data not shown)). A pyridine in agonist 3 is substituted to the carboxylate, and the nitrogen of the pyridine could be engaged in a pyridinium thiolate salt bridge with the side chain of Cys237 (the distance between the nitrogen and the sulfur is only 3.5 A in our model). This interaction is comparable in strength to the carboxylate/arginines interaction of all other RAR ligands. Arg278 may also interact with

the pyridine, according to our docking simulation (Figure 3 C). The development of molecules which conserve this original pyridine, but with improved affinity for the receptor could represent a new generation of modulators of RAR activity. The energy of complexation could be improved for instance by reducing the flexibility and increasing the hydrophobicity of the linker region, which bridges the pyridine to the penta-methyl diphenyl ketone moiety. According to our model, the nitrogens of the linker are not making stabilizing interactions with the receptor and could be replaced by carbons, while one of the two carbonyl oxygens of this domain could be engaged in a hydrogen bond with Ser234 (however, this interaction is probably not critical, since agonist 3 activates also RARβ and RARγ, which have an alanine at position 234). The structure of agonist 3 and the model of its interaction with RARα detailed here represent an avenue towards new RAR agonist ligands.

Conclusion

This report details the rapid discovery of RAR agonists with novel structural features, thanks to a powerful high throughput virtual screening approach, and a research strategy where considerations on existing ligands are avoided. One of the molecules presented here constitute a good framework for the development of a novel series of RAR ligands, deprived of carboxylate and very different from all structures described so far. Such ligands could present more favorable specificity and toxicity profiles, and have important applications in cancer therapy.

Acknowlegement

This research was supported by DOD grant DAMD179818133, a Kaplan Comprehensive Cancer Center grant from NYU Medical Center, NIH grant GM5541801 and DOE grant DEFG0296ER62268 to MS and RA, and NIH grant DK16636 and NYS Empire Award C015710 to HHS. We thank Molsoft LLC for making the latest version of the ICM program available for this research project.

Experimental Section

Modeling of RARα ligand binding pocket

The crystal structure of RARγ-LBD complexed to all-trans RA was used as a template[17] and the three residues in the vicinity of the ligand which are not conserved between the two isoforms were mutated accordingly: A234, M272, and A397 were changed to S234, I272 and V397 respectively. The rotation variables of the side chains within 3.5 A of the mutated residues were unfixed and the energy of the system was minimized in the internal coordinate space, according to the ICM method[26].

Docking of AM580 into RARα

The flexible ligand was docked into a combination of five potential map representations of RARα ligand binding pocket, which account for hydrogen bonding, hydrophobicity, electrostatics and entropy parameters[27]. This rapid docking procedure was followed by a more refined energy minimization of the complex, with a full atom representation of the receptor, and flexible receptor side chains, by an ICM stochastic global optimization algorithm[28,29] as implemented in the Molsoft ICM2.7 program[16].

Virtual screening of the compound structure database

The procedure followed was the same as previously described[8]: each flexible ligand of the Available Chemicals Directory (MDL Information Systems, San Leandro) was docked automatically into the combination of potential maps described above, and assigned a score according to its fit with the receptor. The scoring function included continuum as well as discreet electrostatics, hydrophobicity and entropy parameters[19]. The screening of the over 150,000 ligands database took less than a month on 10 "194 MHZ IP25" processors. The 5364 compounds which scored better (i.e. lower) than -32 were pre-selected for a second round of selection, and were automatically docked into a full atom representation of the receptor, with flexible receptor side chains, according to a global energy optimization in internal coordinates[27,16]. The binding energy of the compounds was then evaluated[30] and the 300 compounds showing the lowest binding energy were selected for further examination. After

careful visual inspection for shape complementarity and hydrogen bonding network, 30 molecules were selected and purchased to be experimentally tested *in vitro*.

Biological activity of the ligand candidates

HeLa cells were transfected by calcium phosphate precipitation using 1 μg of the Gal4-responsive CAT reporter pMC110 and 1 μg of Gal4-hRARα-LBD or 1 μg of Gal4-mRXRβ–LBD. Studies were also performed with the three wild type hRAR isoforms using a ΔMTV-IR-CAT reporter as previously described[20,21]. Cell cultures were supplemented with indicated ligands immediately after addition of the calcium phosphate/DNA precipitate. Media and ligands were replaced after 24 h and cells were harvested and assayed for CAT activity 24 h later.

References

(1) Fitzgerald P, Teng M, Chandraratna RA, Heyman RA, Allegretto E A Retinoic acid receptor alpha expression correlates with retinoid-induced growth inhibition of human breast cancer cells regardless of estrogen receptor status. *Cancer Res* **1997**, *57*, 2642-5

(2) Fanjul AN, Piedrafita FJ, Al-Shamma H, Pfahl M Apoptosis induction and potent antiestrogen receptor-negative breast cancer activity in vivo by a retinoid antagonist *Cancer Res* **1998** *58*: 4607-10

(3) Lu XP, Fanjul A, Picard N, Shroot B, Pfahl M A selective retinoid with high activity against an androgen-resistant prostate cancer cell type *Int J Cancer* **1999** *80*:272-8

(4) Sun SY, Yue P, Dawson MI, Shroot B, Michel S, Lamph WW, Heyman RA, Teng M, Chandraratna RA, Shudo K, Hong WK, Lotan R Differential effects of synthetic nuclear retinoid receptor-selective retinoids on the growth of human non-small cell lung carcinoma cells *Cancer Res* **1997** *57*:4931-9

(5) Sun SY, Yue P, Wu GS, El-Deiry WS, Shroot B, Hong WK, Lotan R Implication of p53 in growth arrest and apoptosis induced by the synthetic retinoid CD437 in human lung cancer cells

Cancer Res **1999** *59*:2829-33

(6) Mologni L, Ponzanelli I, Bresciani F, Sardiello G, Bergamaschi D, Gianni M, Reichert U, Rambaldi A, Terao M, Garattini E The novel synthetic retinoid 6-[3-adamantyl-4-hydroxyphenyl]-2-naphthalene carboxylic acid (CD437) causes apoptosis in acute promyelocytic leukemia cells through rapid activation of caspases. *Blood* **1999** *93*:1045-61

(7) Chandraratna RA Future trends: a new generation of retinoids. *J Am Acad Dermatol* **1998** Oct;*39*(4 Pt 2):S149-52

(8) Schapira M, Raaka B., Samuels HH, and Abagyan R, Rational discovery of nuclear receptor antagonists *Proc Natl Acad Sci U S A.* **2000** Feb 1;*97*(3):1008-13.

(9) Lippman SM, Benner SE, Hong WK Retinoid chemoprevention studies in upper aerodigestive tract and lung carcinogenesis. *Cancer Res* **1994**, *54*, 2025s-2028s

(10) Costa A, De Palo G, Decensi A, Formelli F, Chiesa F, Nava M, Camerini T, Marubini E, Veronesi U. Retinoids in cancer chemoprevention. Clinical trials with the synthetic analogue fenretinide. *Ann N Y Acad Sci* **1995**, *768*:148-62

(11) Liu Y, Lee MO, Wang HG, Li Y, Hashimoto Y, Klaus M, Reed JC, Zhang X Retinoic acid receptor beta mediates the growth-inhibitory effect of retinoic acid by promoting apoptosis in human breast cancer cells. *Mol Cell Biol* **1996** Mar;*16*(3):1138-49

(12) Widschwendter M, Berger J, Daxenbichler G, Muller-Holzner E, Widschwendter A, Mayr A, Marth C, Zeimet AG Loss of retinoic acid receptor beta expression in breast cancer and morphologically normal adjacent tissue but not in the normal breast tissue distant from the cancer. *Cancer Res* **1997** Oct 1;*57*(19):4158-61

(13) Li Y, Dawson MI, Agadir A, Lee MO, Jong L, Hobbs PD, Zhang XK Regulation of RARβ expression by RAR- and RXR-selective retinoids in human lung cancer cell lines: effect on growth inhibition and apoptosis induction. *Int J Cancer* **1998** Jan 5;*75*(1):88-95

(14) Xu XC, Liu X, Tahara E, Lippman SM, Lotan R Expression and up-regulation of retinoic acid receptor-beta is associated with retinoid sensitivity and colony formation in esophageal

cancer cell lines *Cancer Res* **1999** May 15;*59*(10):2477-83

(15) Lin B, Chen Gq, Xiao D, Kolluri SK, Cao X, Su H, Zhang Xk Orphan receptor COUP-TF is required for induction of retinoic acid receptor beta, growth inhibition, and apoptosis by retinoic acid in cancer cells. *Mol Cell Biol* **2000** Feb;*20*(3):957-70

(16) Molsoft, LLC. **1998**. ICM 2.7 program manual

(17) Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D Crystal structure of the RARγ ligand-binding domain bound to all-trans retinoic acid *Nature* **1995** *378*:681-9

(18) Delescluse C, Cavey MT, Martin B, Bernard BA, Reichert U, Maignan J, Darmon M, Shroot B Selective high affinity retinoic acid receptor alpha or beta-gamma ligands. *Mol Pharmacol* **1991** Oct;*40*:556-62

(19) Totrov M., Abagyan R. 1999 *Proceedings of the third annual international conference on computational molecular biology* April **1999**, Lyon, France, ACM Press

(20) Qi JS, Desai-Yajnik V, Greene ME, Raaka BM, Samuels HH. *Mol. Cell. Biol.* **1995** *15*:1817-2

(21) Li D., Desai-Yajnik V., Lo E., Schapira M., Abagyan R., Samuels HH. *Mol Cell Biol* **1999**;*19*(10):7191-7202

(22) Kagechika H, Kawachi E, Hashimoto Y, Shudo K. Retinobenzoic acids. 2. Structure-activity relationships of chalcone-4-carboxylic acids and flavone-4'-carboxylic acids. *J Med Chem* **1989** Apr;*32*(4):834-40

(23) Bernard BA, Bernardon JM, Delescluse C, Martin B, Lenoir MC, Maignan J, Charpentier B, Pilgrim WR, Reichert U, Shroot B Identification of synthetic retinoids with selectivity for human nuclear retinoic acid receptor *gamma Biochem Biophys Res Commun* **1992** *186*:977-83

(24) Dhar A, Liu S, Klucik J, Berlin KD, Madler MM, Lu S, Ivey RT, Zacheis D, Brown CW, Nelson EC, Birckbichler PJ, Benbrook DM Synthesis, structure-activity relationships, and

RARγ-ligand interactions of nitrogen heteroarotinoids *J Med Chem* **1999** Sep 9;*42*(18):3602-14

(25) Benbrook DM, Subramanian S, Gale JB, Liu S, Brown CW, Boehm MF, Berlin KD Synthesis and characterization of heteroarotinoids demonstrate structure specificity relationships. *J Med Chem* **1998** Sep 10;*41*(19):3753-7

(26) Abagyan, R.A. Protein structure prediction by global energy optimization, in *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Volume 3, Eds.**1997** van Gunsteren, W.F. et al., ESCOM Science Publishers BV, Leiden.

(27) Totrov M, Abagyan R Flexible protein-ligand docking by global energy optimization in internal coordinates *Proteins* **1997** Suppl **1**:215-20

(28) Abagyan R, Totrov M Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins *J Mol Biol* **1994**a *235*:983-1002

(29) Abagyan, R.A., Totrov, M.M. & Kuznetsov, D.N. ICM - a new method for protein modelling and design. Applications to docking and structure prediction from the distorted native conformation. *J.Comp.Chem.* **1994**b*15*, 488-506

(30) Schapira M, Totrov M, Abagyan R Prediction of the binding energy for small molecules, peptides and proteins *J Mol Recognit* **1999** *12*:177-190

Legends to Figures:

Figure 1: Docking of a known RARα specific agonist. The RARα selective agonist Am580 was docked into the modeled ligand binding pocket of RARα. A: The complexed ligand (white sticks) superimposes with the crystal structure of bound all-trans RA (green). Hydrogens are not shown for clarity. B: Am580 (CPK display) fits tightly into the receptor's pocket (yellow wire), but for a ketone oxygen, which shares an hydrogen with Ser234 of the receptor (displayed as stick). The receptor in the vicinity of the ligand is shown as a white ribbon. Carbons, hydrogens, oxygens and nitrogen are colored white, grey, red and blue respectively. (Image generated with

Molsoft ICM)

Figure 2: In vitro activity of the novel RAR agonists. HeLa cells were transfected with a Gal4-hRARα-LBD expression vector and a Gal4-CAT reporter gene. The CAT activity induced by each ligand at 0.2, 1 and 10 μM was measured. The activity induced by the natural hormone all-trans RA at 100 nM was used as a positive control. Similar activity was observed with the other two RAR isoforms.

Figure 3: Structure of the novel RAR agonists. Agonists 1, 2 and 3 are shown (respectively A, B and C). Left: chemical structure of the compounds. Right: Representation of the compounds docked into the binding pocket of RAR (important residues are displayed as sticks. A: R274, R278, S289, S234. B: R274, R278, S289. C: C237, R278, S234), and superimposed with the crystal structure of all-trans RA (green). The receptor is represented as a white ribbon. Hydrogens are not displayed for clarity. Color coding: carbons, oxygens, nitrogens, sulfurs, fluorides and hydrogens are colored white, red, blue, yellow, magenta and grey respectively.

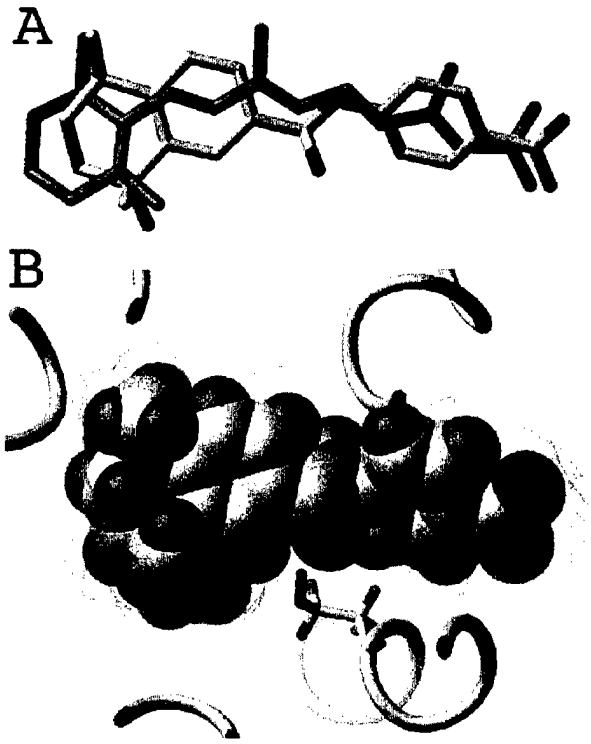


Figure 1

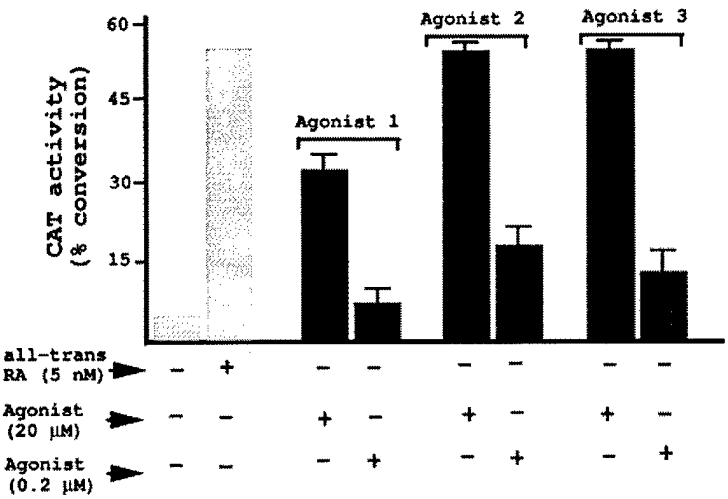

CAT activity
(% conversion)
60
45
30
15
Agonist 1
Agonist 2
Agonist 3
all-trans
RA (5 nM)
Agonist
(20 µM)
Agonist
(0.2 µM)
– + – – – – – –
– – + – + – + –
– – – + – + – +

Figure 2

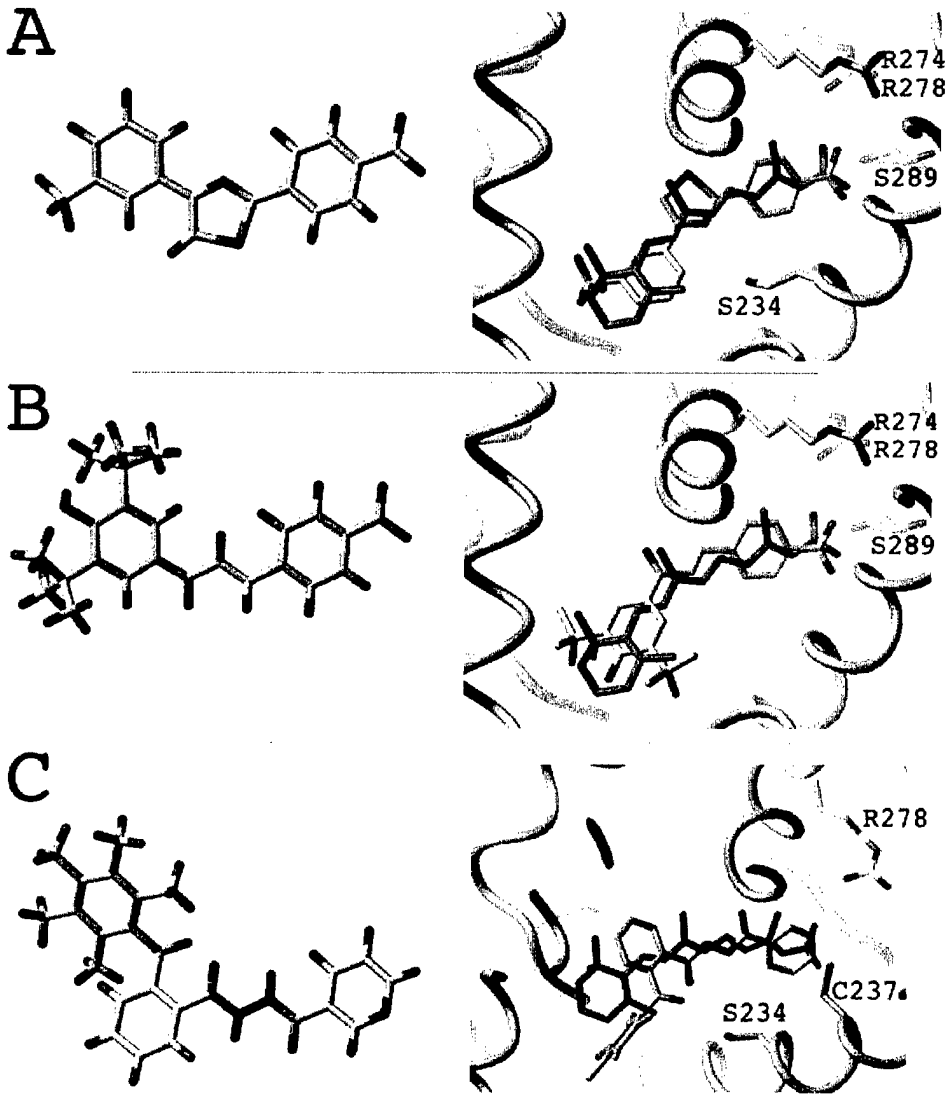


Figure 3