[1] Identification of Novel Nuclear Hormone Receptor Ligands by Activity-Guided Purification

By FELIX GRÜN and BRUCE BLUMBERG

Introduction

Members of the nuclear hormone receptor superfamily share a common architecture typically including a highly conserved DNA-binding domain (DBD) and a ligand-binding domain (LBD). Many of these ligand-modulated transcription factors have specific endogenous ligands and act as ligand sensors that regulate gene expression during development, cellular differentiation, reproduction, and lipid homeostasis. Those receptors lacking known endogenous ligands are termed as "orphan receptors." The identification of natural ligands for this class of nuclear receptors is an important goal in understanding their biology. Much progress has been made in recent years identifying ligands for orphan receptors (reviewed in Refs. 1–4). It is thought that nearly all nuclear receptors evolved from an ancestral estrogen receptor.[5,6] Based on this observation and the existence of conserved LBD sequences, it is not unreasonable to hypothesize that many orphan receptors are ligand-dependent. If many or most orphan receptors do indeed have endogenous ligands, the immediate question that arises is why these ligands have not yet been identified.

Several potential contributing factors for the slow pace of identification of natural ligands can be considered. One is that many previous screens of natural or synthetic ligands are inherently biased toward known bioactive

[1] T. T. Lu, J. J. Repa, and D. J. Mangelsdorf, Orphan nuclear receptors as eLiXiRs and FiXeRs of sterol metabolism, *J. Biol. Chem.* **17**, 17 (2001).

[2] J. J. Repa and D. J. Mangelsdorf, The role of orphan nuclear receptors in the regulation of cholesterol homeostasis, *Annu. Rev. Cell Dev. Biol.* **16**, 459–481 (2000).

[3] T. M. Willson, S. A. Jones, J. T. Moore, and S. A. Kliewer, Chemical genomics: functional analysis of orphan nuclear receptors in the regulation of bile acid metabolism, *Med. Res. Rev.* **21**(6), 513–522 (2001).

[4] W. Xie and R. M. Evans, Orphan nuclear receptors: the exotics of xenobiotics, *J. Biol. Chem.* **276**(41), 37739–37742 (2001).

[5] J. W. Thornton, Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions, *Proc. Natl. Acad. Sci. USA* **98**(10), 5671–5676 (2001).

[6] J. W. Thornton and R. DeSalle, A new method to localize and test the significance of incongruence: detecting domain shuffling in the nuclear receptor superfamily, *Syst. Biol.* **49**(2), 183–201 (2000).

compounds or derivatives thereof and are thus limited in their scope. Ligands or their immediate metabolic precursors may be constitutively present within a cell and have fast turnover rates within narrow concentration ranges. Perturbation through exogenous application of test ligands may therefore be difficult without rigorous control over experimental conditions. Ligand synthesis may be regulated in a very restricted manner in time and space during development. Use of an inappropriate source material may therefore preclude successful purification.

The strategy that we have applied successfully in identifying novel nuclear receptor ligands is based on the key principle of activity-guided purification followed by structure determination. Prior knowledge of an activity's relation to known nuclear receptor ligands is therefore not pertinent. Following sample preparation, extracts are fractionated by suitable HPLC methods and tested for their ability to modulate transcriptional activity of specific nuclear receptors. Central to the success of this approach is a receptor bioassay that has the appropriate sensitivity to detect rapidly ligand-dependent transactivation in small-scale formats. A particular benefit of the modular structure of nuclear receptors is the ability to construct chimeras that retain the specific ligand-dependent nature of a receptor, but facilitate high-throughput screening by utilizing a common reporter construct. Chimeras between the yeast GAL4 DBD and nuclear receptor LBDs have proven invaluable for developing such assays. Reporter constructs exhibit exquisite sensitivity to added ligand in the presence of the chimeric receptors and are independent of endogenous receptor expression making them excellent tools for transactivation screens. Depending on the specific receptor used, quantitative determination of ligand concentrations between 10^{-4} and 10^{-11} M can be measured in tissue culture assays. Therefore, the activity profile of a specific chimeric receptor in response to ligand mixtures (e.g., HPLC fractions) allows the candidate ligand to be purified to homogeneity and it is structure determined in the absence of prior information about its structure.

Although the information presented below is drawn primarily from our experiences of novel retinoid receptor ligand purification, the chapter is organized with the goal of presenting general strategies and protocols that are expected to be widely applicable for lipophilic ligands isolated from both tissue sources or dilute xenobiotic ligands present as environmental endocrine disruptors. Special focus is given to sample preparation and extraction techniques, HPLC fractionation methods, and the receptor activation transcription assay. A brief outline of mass spectrometric approaches to structure determination is also included as an introduction on how to proceed when an activity has been purified and to give a sense of the scale-in material required.

Sample Preparation and Extraction Methods

The methods described below will use retinoids as example compounds. It is important to note, however, that the approach is rather general and can readily be adapted for other types of compounds. Extracting and identifying compounds that activate retinoid receptors is straightforward, although not trivial.[7] The primary considerations are that one must identify an appropriate source of material and then extract undegraded and unmodified candidate compounds that are then fractionated and tested for retinoid activity. In the case of embryos, one typically pools embryos from stages wherein the receptor is expressed. Tissues expressing the receptors of interest are also good candidates. Lastly, environmental samples such as water or sediments may be used as sources of potentially unknown retinoids.

A good starting point is to utilize a method that recovers the unknown compound quantitatively from the source material, while maintaining its stability, chemical form, and biological activity. This extract should be tested directly for biological activity and also crudely fractionated using semi-preparative scale HPLC and the fractions tested for activity. It is not always the case that activity can be detected in the crude extract; however, if no activity is detected in the fractions, then the extraction method must be changed or optimized. This step ensures that (1) there actually is an activity to purify and (2) a reference method exists that can be used to validate subsequent preparative methods such as solid-phase extraction or solvent partition. It is important to emphasize that the activity must be successfully identified at this stage before proceeding to large-scale purification.

General Considerations

For ligand extracts prepared from tissue samples, or that require special consideration in terms of sensitivity to light, temperature, or biochemical reactivity, we routinely use a homogenous liquid–liquid technique.[7,8] The method is both rapid and gentle requiring no prolonged extraction steps, elevated temperatures, or pH extremes. It is particularly well suited to protein-rich samples since deproteination occurs during extraction and the extracts are compatible with direct HPLC injection. Extraction volumes are kept relatively small and this method can be easily scaled to accommodate

[7] B. Blumberg, J. Bolado, Jr., F. Derguini, A. G. Craig, T. A. Moreno, D. Charkravarti, R. A. Heyman, J. Buck, and R. M. Evans, Novel RAR ligands in Xenopus embryos, *Proc. Natl. Acad. Sci. (USA)* **93**, 4873–4878 (1996).

[8] S. W. McClean, M. E. Rudel, E. G. Gross, J. J. DiGiovani, and G. L. Pede, Liquid-chromatographic assay for retinol (vitamin A) and retinol analogs in the therapeutic trials, *Clin. Chem.* **28**, 693–696 (1982).

larger volumes up to several liters. The sample is first homogenized with 0.4 volumes of acetonitrile:*n*-butanol (1:1) using a polytron, Dounce, or other appropriate homogenizer. Homogenizing for 1 min using the polytron at maximum speed is adequate. Aqueous samples, such as serum, can be vortexed vigorously for several minutes. Phase separation is accomplished by addition of 0.3 volumes of saturated dibasic potassium phosphate solution. The samples are homogenized or vortexed for an additional minute and then centrifuged at $10{,}000 \times g$ for 1–10 min at 4°C to separate the layers. After centrifugation, the organic phase is removed for HPLC separation. Proteins form a gelatinous phase-lock between the aqueous (bottom) and the organic (top) layers and are thereby effectively removed. Addition of the antioxidant *tert*-butylated hydroxytoluene (BHT) at 1 μM during the extraction is recommended to reduce nonspecific oxidation of the sample.

Solutions
Acetonitrile:1-butanol (1:1 v/v)
Saturated K_2HPO_4 solution, pH 7.5
1 m*M* BHT in methanol
Phosphate-buffered saline (PBS)

Example for small sample volumes (*Eppendorf tube scale*)
0.8 ml Aqueous samples
0.32 ml Acetonitrile:1-butanol (1:1)
0.28 ml K_2HPO_4
1.4 ml Total volume

Centrifuge at 14,000 rpm in a benchtop microcentrifuge for 1 min. Approximately, 200 μl of organic extract is obtained that can be injected directly for HPLC analysis without further sample cleanup.

Example for embryos, tissues or solid material
20 ml Homogenate
8 ml Acetonitrile:1-butanol
6 ml Saturated K_2HPO_4
34 ml Total

For embryos, tissues, and freeze-dried samples, the material should be homogenized for at least 1 min in a 10-fold excess of PBS before addition of acetonitrile–butanol. After addition of the saturated potassium phosphate, vortex, the material is transferred to polypropylene or polyallomer high-speed centrifuge tubes and it is centrifuged at $10{,}000 \times g$ for 10 min at 4°C in a Sorval SA-600 rotor or similar rotor. The top organic layer is transferred

to a new tube and centrifugation is repeated to remove any carryover of particulate matter. The recovered material can be concentrated by rotary evaporation and reconstituted in acetonitrile–butanol or simply injected onto the HPLC column in appropriately sized aliquots (see below).

Alternative Organic Extraction Methods

Acetonitrile–butanol extraction is convenient and usually preferred for compounds of unknown chemical properties or volatility. It is not the best method for extracting acids or if the material is known to be stable and not volatile. Eichele and co-workers have published extensively on the extraction and characterization of retinoic acids (reviewed in Wedden *et al.*[9]). In addition, there are a variety of other types of liquid–liquid organic extractions, e.g., the Folch method (methanol–chloroform 2:1).[10] It should be noted that the use of chloroform is best avoided due to its toxicity and frequent acidity. Dichloromethane gives similar extractions and is not acidic.

Solid-Phase Extraction

Materials

500 g Amberlite XAD-2 resin (Rohm & Haas)
Large Soxhlet extractor
Rotary evaporator
Water filtration canisters (paper or glass fiber prefilter, connectors)
Pond pump
Large Buchner funnel
Acid-washed detergent-free glassware
4 liter Methanol
4 liter Acetone
4 liter Hexane
4 liter Dichloromethane
30 liter Double-distilled deionized water
20 cm Glass fiber filters
Glass wool

[9] S. Wedden, C. Thaller, and G. Eichele, Targeted slow-release of retinoids into chick embryos, *Methods Enzymol.* **190**, 201–209 (1990).

[10] J. Folch, M. Lees, and G. H. S. Stanley, A simple method for the isolation and purification of total lipids from animal tissues, *J. Biol. Chem.* **226**, 497–509 (1957).

For preparative extractions of large aqueous samples, e.g., serum or environmental water samples, liquid–liquid extraction becomes impractical due to the large volumes of organic solvents required. A suitable solid-phase extraction method with broad affinity for hydrophobic organic compounds (HOC) is therefore desirable. We have successfully employed XAD-2 resin beads (Rohm & Haas) to capture retinoid receptor activators from environmental water samples. Alternatively, other high-flow matrices (e.g., HP-20, XAD-7, C18, silica) with affinity for the receptor activity in question may be substituted. It may be necessary to try several matrices before finding one that is suitable XAD-2 resin as supplied by the manufacturer requires extensive cleanup prior to use in order to remove residual synthetic reaction products. The Environmental Protection Agency (EPA) has developed and standardized methods for the preparation and use of solid-phase resins such as XAD-2,[11] and we have adopted these without change. Briefly, resin is extracted sequentially with methanol, acetone, hexane, and methylene chloride for 24 hr each in a Soxhlet extractor, followed by the reverse order of solvents for sequential 4 hr extractions to bring the resin back to a polar solvent miscible with water. The methanol is replaced by washing the beads six times with 10 volumes of double-distilled water. Beads are stored under water in EPA-certified glass sample bottles until use. Storage should be for 3 months or less and care should be taken to avoid mechanical damage during column packing. The final 4-hr hexane extract is used as an XAD-2 blank.

XAD-2 may be used for small-scale as well as large-scale extractions. Embryos or tissues to be used for solid-phase extraction are homogenized in deionized water or phosphate buffer at approximately 50 mg/ml and the extracts are clarified by centrifugation or filtration. The supernatant is extracted by stirring with preconditioned XAD-2 at a ratio of 1 volume of resin slurry to 5 volumes of supernatant. The mixture is stirred at room temperature for 4 hr with a propeller stirrer, and the beads recovered by decanting the supernatant and removing as much liquid as possible by vacuum filtration. The resin is rinsed with 10 volumes of deionized water and the liquid is again removed by filtration. Adsorbed material is recovered by stirring with 10 volumes of methanol for 2 hr, and the methanol recovered as above. The process is repeated with 10 volumes of acetone and the acetone recovered. Both methanol and acetone soluble materials are combined, clarified by filtration, rotary evaporated, and reconstituted in dimethyl sulfoxide or chloroform–methanol (1:2 ratio) under argon.

[11] E. Crecelius and L. Lefkovitz, HOC Sampling Media Preparation and Handling; XAD-2 Resin and GF/F Filters, Standard Operating Procedure MSL-M-090-00 ed. US EPA, 1994.

For very large-scale preparative purposes (e.g., environmental samples), XAD-2 resin is packed into polycarbonate water filtration canisters modified to contain resin and is fitted with 3/4-inch connectors. Glass wool is packed into the central perforated drainage tube to prevent resin from being flushed out. The inverted canister is partially filled with water and a slurry of resin poured around the central drain tube. Excess water is aspirated from the center taking care to avoid introduction of air bubbles that can lead to channeling and a reduction in the exposure of beads during pumping. Glass wool is packed on top of the resin bed and the canister base with O-ring is screwed tight. Upright packed canisters should be gently flushed with clean deionized water to verify that the canister does not leak water from the seals or resin from the drain outlet. Canisters premodified for chemical media may be obtained from aquarium suppliers (e.g., Filtronics, Oxnard CA), although these have somewhat lower capacity. Three canisters each containing approximately 500 ml of packed resin can be connected in series to retain a suitable flow rate of 2–3 liter/min. An additional canister containing a paper or a glass fiber prefilter is connected between the pump and the first canister to remove particulates when collecting environmental water samples. The whole assembly is connected by 3/4-inch polyethylene tubing and secured using metal screw collar retainers. A typical pond or aquarium circulating pump (e.g., Eheim, Little Giant, Iwaki, Hydrothruster) with a magnetic drive is utilized to provide the flow. The canister valves are set to provide a steady flow of 2–3 liter/min, which is suitable to prevent undue compression of the resin bed. After several hours, the flow rate should be checked and the prefilter exchanged if fouling is a serious problem. The intake hose is placed within a weighted bucket on the lakebed to provide protection from sediment disturbances, weeds, and other larger debris during setup. The columns are run for 24–48 hr to allow the XAD-2 resin to scavenge HOCs from the water source. Columns are then sealed and shipped to the lab for resin processing.

Resin slurry (500 ml per extraction) is transferred to a 4-liter glass beaker for extraction. Excess water is removed by aspiration. The resin is then sequentially extracted three times with 1 liter each of methanol, acetone, and methylene chloride for 1 hr at room temperature with gentle stirring using a propeller stirrer (Fisher), filtered through a glass fiber filter on a large Buchner funnel and concentrated on a rotary evaporator. The concentrated extract is redissolved in a minimum of organic solvent, e.g., methanol and adjusted to approximately 50–75% solvent/25–50% aqueous buffer depending on the point at which precipitation of solutes is noticeable. The extract is cleared by centrifugation. Any precipitate is back-extracted with 75% methanol/25% buffer. Remaining insoluble material should be discarded.

The pooled organic phases are loaded in aliquots onto a reversed-phase C18 column equilibrated in buffer (typically 50 m*M* ammonium acetate, pH 6.5) until the entire extract is on the column. Subsequently, a linear gradient of buffer–methanol–chloroform is employed at a flow rate appropriate for the size of column used and fractions are collected. Under these conditions, the materials do not begin to elute until the organic component of the mobile phase reaches a critical level for each type of compound. In this way, handling time is minimized and the requirement for drying down large volumes of organic extracts is eliminated.

Solvent Partition as a Prepurification Step

Once a sample is known to contain activity, it is often valuable to determine whether it can be partially purified by solvent partition prior to chromatography. This enables one to reduce the complexity of a mixture, which can reduce greatly the time spent in optimizing later HPLC separations. Solvent partition employs sequential extractions between immiscible solvents of different chemical properties. One begins with a known amount of retinoid activity then adds the sample to a large excess of methanol (> 20 volumes) in a separatory funnel. An equal volume of iso-octane or *n*-heptane (*n*-heptane is easier to work with) is then added and the mixture is shaken vigorously. After phase separation, the lower methanol phase is removed and reextracted with another aliquot of *n*-heptane, shaken as above, and the phases separated. After three extractions, the *n*-heptane and methanol phases are rotary evaporated to dryness, reconstituted, and tested for activity. If the activity partitions into the nonpolar phase it is used directly for HPLC purification. If it partitions into the polar methanol phase, the next step is to partition the activity between ethyl acetate and water using the same procedure. If the activity partitions into the ethyl acetate phase it is then further purified by HPLC. If it partitions into the aqueous phase, further partitioning between water and *n*-butanol is conducted and the resulting extracts tested for activity. In testing each of these partition steps, it is important to quantitate carefully the fraction of the active component recovered. If it is not recovered quantitatively then solvent partition is unsuitable as a prepurification step.

Reversed-Phase HPLC

We utilize a Waters 600 Delta HPLC system running Millennium 32 chromatography manager software. The key components of the system are dual Rheodyne 7725i injection valves for semi-prep and analytical injection switching, a Waters 600E pump and 996 photodiode array detector, and

a Pharmacia SuperFrac fraction collector. We favor Vydac 218TP (fully endcapped) or 201TP (non-endcapped) C18 columns for reversed-phase chromatography although the quality of modern columns makes the exact choice of manufacturer and chemistry a personal preference. Guard columns should be used throughout to protect and prolong the lifetime of the columns. A good set of columns for extract purification comprises a 25×250 mm preparative, 10×250 mm semi-preparative and 4.6×250 mm analytical columns. A set of columns and injection syringes dedicated for ligand purification is well worth the investment to avoid contamination from standards. In all cases, PDA monitoring of a wide spectrum of wavelengths at one spectrum per second is employed. When methanol is used as the solvent, we typically scan a window of 220–600 nm. Acetonitrile is optically clear at low wavelengths, therefore a window of 190–600 nm is used.

Crude extracts are first fractionated on the preparative column. The column is preequilibrated by executing mock runs of the gradient elution method 2–3 times until the baseline is stable and clear of contaminant peaks. It is important to scrupulously clean the injector ports between runs, especially if the HPLC system has been used previously for analysis of standards. We always perform a blank run prior to sample loading to enable collection of solvent only controls for each later fraction. The column is loaded under initial flow conditions, typically 95–100% of an aqueous buffer, e.g., 50 m*M* ammonium acetate, pH 6.5. Multiple aliquots (2 ml or less) are injected every 3 min with baseline monitoring. If breakthrough of injected compounds is observed, the percentage of methanol in the sample or volume of the injected aliquot is reduced. After the final injection, the column is allowed to return to baseline absorbance value before starting the gradient run. An extract from 500 ml packed XAD-2 resin is normally split into 2–3 preparative runs to avoid column overloading and the collected fractions pooled.

Gradient elution, as outlined in Table I, cycles the column from 100%, the initial conditions within 80 min. All of the retained compounds will elute within 60 min. Fractions are collected at 1-min intervals (flow rate 8 ml/min). We typically add BHT to the empty tubes to achieve a final concentration of 1 μM. This addition ensures that BHT is continuously present. After collection, the samples are flushed with argon, stored at -80°C, and protected from light until analyzed in the reporter assay.

Active fractions from the preparative runs are pooled, diluted 3:1 with aqueous buffer, reinjected, and subjected to further fractionation by HPLC. As a second column, we use a Vydac 218TP510 semi-preparative C18 column. The eluting solvent is an acetonitrile gradient from 0 to 100%. One-minute fractions (flow rate 2 ml/min) are collected and 100 μl aliquots are removed for assays. Bona fide activity peaks should

TABLE I
GRADIENT CONDITIONS

Time (min)	Flow (ml/min)	% A	% B	% C	% D	Change
(A) Preparative C-18 column with buffer–methanol–dichloromethane gradient						
0	8	100	0	0	0	–
5	8	30	70	0	0	Linear
35	8	0	100	0	0	Linear
40	8	0	100	0	0	Linear
50	8	0	40	60	0	Linear
60	8	0	100	0	0	Jump
65	8	100	0	0	0	Jump
80	0	100	0	0	0	Jump
(B) Semi-preparative C-18 column with buffer–acetonitrile–dichloromethane gradient						
0	3	100	0	0	0	–
5	3	80	0	0	20	Linear
40	3	0	0	0	100	Linear
42.5	3	0	0	0	100	Linear
50	3	0	0	100	0	Linear
55	3	0	0	0	100	Jump
65	3	100	0	0	0	Jump
80	0	100	0	0	0	Jump
(C) Analytical C-18 column with buffer–mixed methanol/acetonitrile (1:1) gradient						
0	1	100	0	0	0	–
2.5	1	40	30	0	30	Linear
30	1	30	35	0	35	Linear
35	1	0	50	0	50	Linear
40	1	100	0	0	0	Jump
60	0	100	0	0	0	Jump

Eluent	Composition
A	50 m*M* Ammonium acetate, pH 6.5
B	100% Methanol
C	100% Dichloromethane
D	100% Acetonitrile

show consistent patterns of activation between column runs and some indication of dose-dependency when diluted or compared with adjacent fractions.

Additional steps can be performed on high-quality analytical columns from this point onwards as the amount of material will no longer saturate and impede column performance. Choices of column chemistry and mobile phase will be dependent on the specific ligand and need to be addressed on a case-by-case basis. C18 columns are ideally suited for fractionating nonpolar retinoids. Our preferred analytical column is a Vydac 201TP54

developed with a shallow methanol gradient. The slight difference in solvent polarity may result in a change in order of elution of specific coeluting components, which together with careful optimization of gradient conditions is often sufficient to give baseline separation of compounds observed in the PDA UV spectrum. Additional changes in parameters such as buffer pH or counterion and column temperature may be useful. Organic acids, such as retinoic acid, demonstrate a significant shift in retention time (several minutes) when changing from neutral to acidic conditions due to the change in ionization states. Acetic acid (1%) is useful as a relatively gentle modifier and can be removed easily during subsequent sample processing; however, care must be taken to ensure that the fractions retain activity under these conditions. Neutralization with buffer following elution is recommended to prevent losses from acid catalysis.

Caveats

The initial stages of any purification should be undertaken with the utmost care to avoid the degradation and modification of potentially unstable compounds. Such steps include eliminating rotary evaporation where feasible to avoid oxidation and potential loss of volatile compounds, ensuring that contact with oxygen is minimized, the continuous presence of suitable antioxidants (e.g., BHT), and rapid testing of fractions for biological activity. It is often advisable to purchase screw-capped test tubes and blanket the samples with argon prior to freezing them. Neutral buffers without primary amines are favored to minimize chemical modification and degradation. Whenever feasible, all extractions, fractionation, transfections, and other analyses should be performed using subdued lighting. Once some idea of the lability of the compounds under study is obtained, these precautions may be modified accordingly.

How Much Material is Required for Chemical Characterization?

One difficulty often encountered when identifying new receptor ligands is that the activation assays can identify nM or pM levels of compounds, whereas tens of μM or more may be required for chemical analysis. Generally speaking, one requires milligrams of pure compound to characterize an unprecedented carbon skeleton by mass spectrometry, ^{1}H- and ^{13}C-NMR. This requirement may be reduced if an NMR with micro- or nanoinverse detection probes is available. For example, we required less than 20 μg of pure material to characterize 3-hydroxyl ethyl benzoate glucosamine by ^{1}H-NMR and tandem mass

spectrometry.[12] Typically, tens of micrograms are required to identify a known compound using mass spectrometry.

Activity-Guided Fractionation

One of the most important components of our ligand identification method is the activity-guided fractionation. The purification is inherently unbiased with respect to any *a priori* knowledge of the compound's relation to known ligands since it is guided solely by following the compound's potential to transactivate the receptor LBD. Each fraction is tested for activity in the cotransfection assay and active fractions selected for further study. The identity of the compound is determined after it is purified to homogeneity through chemical characterization. This approach enables the identification of compounds with unusual fractionation or absorption properties. To date, we have identified and characterized novel ligands for the retinoic acid receptor,[7] and identified the orphan receptor BXR as a specific receptor for benzoates[12] using activity-guided fractionation.

For high-throughput screening of fractions, we use a calcium phosphate/DNA precipitate protocol adapted for a 96-well format. This low-cost transfection method saves on reagents, and when combined with multichannel dispensers and plate readers, allows a single operator to assay easily as many as several thousand data points per experiment.

Success in ligand purification is determined in large part by the sensitivity and robustness of the reporter assay. A key requirement is that the sensitivity of the assay must be comparable to the concentration of the active component in the material being studied. For receptors that have high-affinity ($\sim$nM K_d) ligands such as RAR, assay sensitivities in the subnanomolar range are optimal and readily attained (Fig. 1). This sensitivity allows one to detect even weak activation and to minimize the use of precious material. To achieve such sensitivity, considerable attention to the optimization of assay conditions is essential. Particularly important is to optimize the transfection efficiency, which is strongly influenced by the pH of the BES phosphate buffer, incubator CO_2 levels, and buffering capacity of the growth medium. For this reason, empirical determination of optimal conditions using the conditions below as a starting point of reference should be conducted. Consistent results can be attained but require rigorous

[12] B. Blumberg, H. Kang, J. Bolado, Jr., H. Chen, A. G. Craig, T. A. Moreno, K. Umesono, T. Perlmann, E. M. De Robertis, and R. M. Evans, BXR, an embryonic orphan nuclear receptor activated by a novel class of endogenous benzoate metabolites, *Genes Dev.* **12**(9), 1269–1277 (1998).

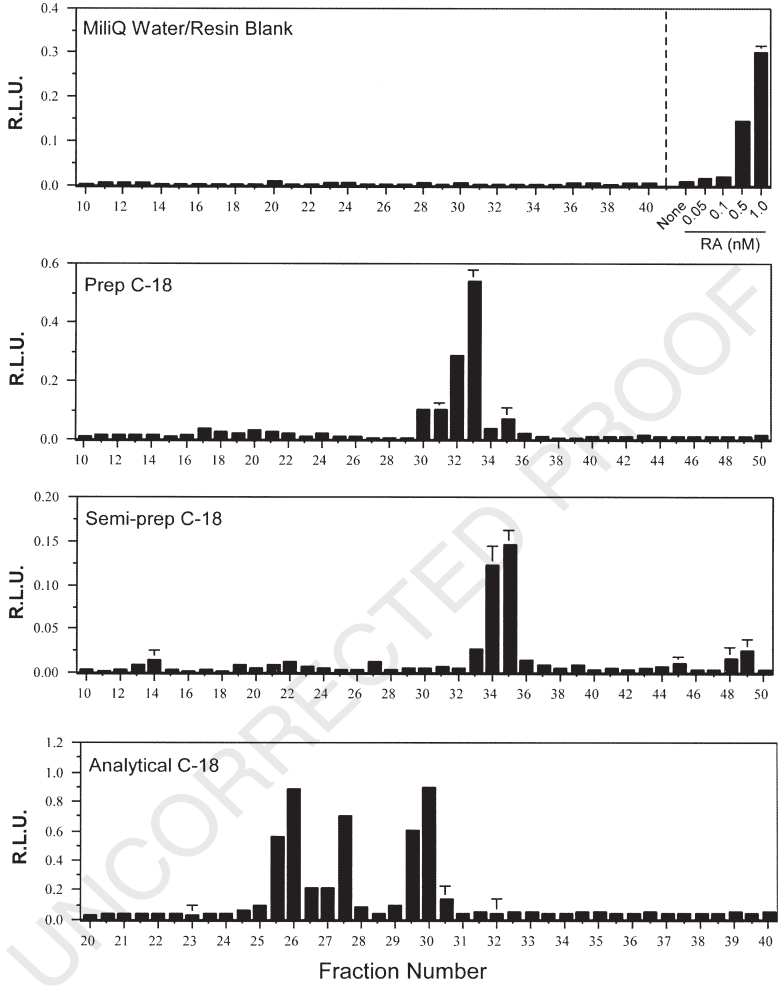

FIG. 1. Activity-guided fractionation of Minnesota lake water. HOCs from a Minnesota lake with a high incidence of malformed amphibians or control Milli-Q water were captured by solid-phase XAD-2 resin extraction, fractionated by C18 reverse-phase HPLC, and tested in receptor activation assays. Aliquots from 1 min fractions were tested for GAL4-RARα-mediated transactivation of luciferase reporter constructs in transient transfection assays and compared with all-trans retinoic acid standards. An environmental retinoid activity is detected in fractions 30–33 of the preparative C18 column. Fractions 32 and 33 were pooled and rechromatographed on the semi-preparative C18 column. Fractions 34 and 35 were pooled and rechromatographed on the analytical C18 column, where the activity resolved into three distinct peaks. No activity is seen in the corresponding water and resin controls (top); retinoic acid standards could be detected in the subnanomolar range (top). Bars represent the means $\pm$ S.E.M. of triplicates normalized to β-galactosidase controls and expressed as relative luciferase units (R.L.U.).

adherence to a laboratory standard experimental procedure. Individual operators will frequently obtain suboptimal performance even with the same reagents.

Calcium chloride solution and BES buffer are prepared as 500 ml 2× stocks. The optimal pH of the BES phosphate buffer stock should be determined for the specific conditions, i.e., cell line, medium, and incubation times, used for the ligand screen. This optimization is done most easily by test transfections with CMX-β-galactosidase expression plasmid after incremental addition of 5 μl aliquots of 0.5 N NaOH or HCl to 1 ml samples of the 2× BES buffer. The 500 ml stock buffer is subsequently adjusted volumetrically based on the sample that the gives best results. The buffer is stable at room temperature for 6 months or can be frozen as 10 ml aliquots at −20°C.

2× BES pH 6.95 buffer	
NaCl	8.18 g
BES	5.33
150 m*M* Na_2HPO_4	5.0 ml
(pH the solution carefully to 6.95)	
Distilled H_2O to	500 ml
Filter sterilize	

2× $CaCl_2$ solution	
1 *M* Tris, pH 7.5	0.5 ml
0.5 *M* EDTA, pH 8.0	0.1 ml
$CaCl_2 \cdot 2H_2O$	18.37 g
Distilled H_2O to	500 ml
Filter sterilize	

Plasmid DNAs are purified by alkaline lysis followed by double banding in CsCl density gradients and stored at 4°C as concentrated (>1 mg/ml) stocks in TE/10 buffer (10 m*M* Tris, pH 7.5, 0.1 m*M* EDTA). Sufficient plasmid DNA (2–5 mg) should be prepared to allow regular large-scale transfections over several months. Column-based purification methods may be used but these typically cost at least 10-fold more per mg of plasmid obtained than CsCl density-gradient centrifugation.

Transfection Protocol

1. COS-7 cells are grown to 80% confluency in 10-cm tissue culture petri dishes. Cells are removed with 0.01% trypsin/0.3 m*M* EDTA in PBS. Pellet was centrifuged at 1500 × *g* for 5 min.

2. Count and seed 96-well tissue culture plates at 5000 cells/well dispensed in 100 μl 10% FBS/DMEM.
3. Incubate cells for >5 hr prior to transfection to allow cells to reattach. Cells should be at a density of 30–50% confluence.
4. Prepare calcium phosphate/DNA precipitate within the range of 10–20 μg DNA/ml calcium phosphate solution. The plasmid mix will include a GAL4-DBD/receptor-LBD fusion effector plasmid, a luciferase reporter plasmid containing GAL UAS promoter sequences [e.g., tk(MH100)$_4$-luc], a constitutively expressed β-galactosidase reporter control (e.g., CMX-β-gal), and carrier DNA. The optimal ratio of receptor effector plasmid to luciferase reporter should be determined by preliminary titration experiments. For receptors with known ligands, a complete dose–response curve will indicate conditions of maximum sensitivity and fold induction. Frequently, the optimum effector plasmid amount will lie in the range of 50–1000 ng/plate at a ratio between 1:5 and 1:10 of effector–reporter. We use typically a ratio of 1:5:5:4 of receptor–luciferase reporter–β-galactosidase control–carrier DNA, i.e., 1 μg pCMX-GAL4-RARα, 5 μg tk-(MH100)$_4$-luc, 5 μg pCMX-β-gal: 4 μg pBluescript for a total of 15 μg/96-well plate. The plasmid DNAs are mixed into 0.6 ml $2\times$ $CaCl_2$ solution and then added dropwise with vigorous vortexing to 0.6 ml $2\times$ BES buffer. A fine precipitate should form. Allow the precipitate to mature for 5 min, but not longer than 15 min, with occasional vortexing.
5. Mix with 10 ml prewarmed 10% FBS/DMEM medium and dispense 100 μl per well. A fine, even precipitate coating the cells should be visible after several minutes. A large clumpy precipitate or a sparse precipitate is an indication that the transfection conditions are not optimal.
6. Incubate plates for 5–24 hr at 37°C in a humidified incubator at 5% CO_2 in air.
7. After transfection, the medium is aspirated from each well and cells washed gently twice with 200 μl prewarmed PBS to remove precipitate and serum. Medium is replaced with 100 μl/well serum-free ITLB/DMEM or 10% charcoal-stripped FBS/DMEM.
8. Test fractions from HPLC separations are prepared as follows. Aliquots of 100–500 μl from each fraction are transferred to siliconized Eppendorf tubes and rotor evaporated unheated in a SpeedVac. Samples are resuspended in 8 μl of 100% ethanol or DMSO, mixed with 400 μl ITLB/DMEM medium (or other suitable medium), and 100 μl dispensed into each of triplicate wells of the transfected 96-well cell culture plate. When available, a

standard curve of a positive ligand control should be included on every 96-well plate.

9. Incubate plates for a further 24–48 hr at 37°C in a humidified 5% CO_2 incubator to allow for maximum ligand-dependent transcriptional responses. An initial test from 12 to 48 hr is suggested to verify the best incubation times as the final luciferase activity will reflect the integration of multiple factors. These include effects of ligand metabolism (including potential metabolic activation or breakdown), receptor protein levels, decay of basal luciferase levels after switching to minimal or defined medium.
10. Media is removed by careful aspiration and plates washed once with PBS.
11. 150 μl/well complete cell lysis buffer is added and plates shaken on a Titer Plate Shaker for 10–30 min.

Luciferase and β-Galactosidase Assays

Cell lysis buffer

1 *M* Tris–PO_4, pH 7.8	25 ml
Glycerol	150 ml
CHAPS	20 g
Phosphotidyl choline (lecithin from egg-yolk)	10 g
BSA	10 g
Distilled H_2O to	1000 ml

Stir the lecithin on a hot plate to 70°C in 100 ml deionized water until completely dissolved. Add the remaining components except BSA. BSA is dissolved separately in 50 ml distilled water and added when the temperature is below 50°C. Filter when warm through a 0.45 μm filter. Store at 4°C. The lysis solution should be a pale straw yellow color. Add the remaining reagents shortly before use.

Add fresh per 15 ml lysis solution

0.1 *M* EGTA, pH 8.0	600 μl
1 *M* $MgCl_2$	120 μl
1 *M* DTT	15 μl
0.2 *M* PMSF in methanol	30 μl

Luciferase assay reagent (per 10 ml)

200 m*M* Tricine–NaOH, pH 7.8	1 ml
Mg^{2+} stock solution	0.1 ml

0.5 *M* EDTA	2 μl
1 *M* DTT	0.3 ml
2.5 m*M* Coenzyme A	1 ml
20 m*M* ATP	0.5 ml
1 m*M* Luciferin	1 ml
Distilled H_2O to	10 ml

Mg^{2+} stock solution

$(MgCO_3)_4 \cdot Mg(OH)_2 \cdot 5H_2O$	10.394 g
$MgSO_4 \cdot 7H_2O$	13.162 g
Distilled H_2O to	160 ml

Stir until dissolved

25 mM Coenzyme A stock solution
0.01 g in 5 ml H_2O
Dispense into aliquots and store at −20°C

100 mM ATP stock solution
0.11 g in 5 ml H_2O
Dispense into aliquots and store at −20°C

10 mM D-Luciferin in methanol
Protect from light, store at −20°C

β-Galactosidase assay buffer

β-Gal base solution	10 ml
β-Mercaptoethanol	30 μl
O-Nitrophenyl-β-galactopyranoside	10 mg

β-Gal base solution

$Na_2HPO_2 \cdot 12H_2O$	21.49 g
$NaH_2PO_4 \cdot 2H_2O$	6.24 g
KCl	0.75 g
1 *M* $MgCl_2$	1 ml
Distilled H_2O to	1000 ml

Autoclave and store at room temperature

1. For luciferase and β-galactosidase assays, 50 μl lysate aliquots are transferred to solid white luminometer and clear, flat-bottom 96-well plates.
2. Dispense 100 μl/well luciferase assay reagent to the samples in the luminometer plates, mix briefly on a microtiter plate shaker, then

read on a 96-well plate luminometer (e.g., Torcon R-7 or Dynex MLX) set to measure in cycle mode. The stabilized luciferase reaction has a half-life of approximately 5 min permitting quantitation of the entire plate.

3. For β-galactosidase activity determination, dispense 100 μl β-galactosidase assay buffer to each well, mix briefly on a microtiter plate shaker and incubate at 37°C (e.g., in a bacterial incubator). Follow the color development by periodic measurement on a 96-well plate spectrophotometer set at 405 nm. When the OD reaches 0.5–1.0 AU (15–60 min), stop the reaction by addition of 100 μl 1 *M* sodium bicarbonate, mix briefly and remeasure the OD. Note the time.Luciferase and β-galactosidase data are exported to Excel spreadsheet templates for data analysis and graphing. Background values for luciferase and β-galactosidase activity are determined from mock-transfected wells and can be subtracted from the experimental values. Individual wells are expressed as luciferase values normalized to β-galactosidase activity and plotted as Relative Luciferase Units per O.D. 405 nm per minute. Triplicate assays are averaged and plotted as the mean $\pm$ S.E.M. Fold activation is expressed as the activation observed relative to the average for solvent only controls.

Structure Determination by Mass Spectrometry

After sequential rounds of HPLC purification and receptor activity-guided assays, the active fraction(s) should be sufficiently enriched for preliminary structural studies by mass spectrometry. Consultation with an experienced mass spectrometrist is strongly recommended in planning an analytical strategy and in the interpretation of data when attempting to elucidate the chemical structure of an unknown compound. While every situation presents unique challenges, most structure characterizations follow a general scheme as outlined below:

1. Determine complexity of active fraction and identify molecular ion of active component.
2. Obtain exact mass measurement and derive molecular formula.
3. Obtain electron-impact (EI) spectra for structural analysis and library search.
4. Compliment structural analysis with ^{1}H- and ^{13}C-NMR studies where appropriate. Mass spectrometry is unlikely to provide identification of structural isomers unless these can be chromatographically separated and compared with authentic standards.

The two biggest hurdles facing the investigator concern purity and quantity. Some of the most common mass spectrometry techniques outlined below utilize additional chromatographic separation with gradients of either solvent (LC) or temperature (GC). These methods can often provide sufficient resolution to obtain spectra on individual components in mixtures. Purification to homogeneity, as judged for instance by peak symmetry and peak purity plots (UV absorbance ratios) of chromatographic peaks, is therefore not strictly necessary but remains desirable if possible. The main benefit of a homogenous preparation is increased sensitivity. While mass spectrometry can be an exquisitely sensitive methodology for detection, structure determination of unknown compounds requires sufficient material for method development and collection of high-quality spectra. Quantities of 1 nm per compound are usually adequate for gathering basic information, but complete structure determination may well require more, especially if compounds of interest require derivatization prior to analysis.

Exact Mass Measurement by ES-MS

Our strategy initially is to analyze unknown samples by electrospray ionization TOF mass spectrometry (ES-MS) in positive ion mode by flow-injection analysis on a Micromass LCT instrument. The method requires no special sample preparation, is compatible with most HPLC fractionation solvents, can analyze broad classes of organic compounds, and is sensitive in the submicromolar range. Compounds with one or more readily ionizable functional groups (e.g., carboxyl-, amino-, or hydroxyl) should give satisfactory spectra. Samples are prepared in a minimal volume of methanol, typically 100 μl in conical borosilicate glass autosampler tubes, and 50 μl injected directly from a Gilson 231XL autosampler into the methanol solvent stream (200 μl/min). Mass spectra between 100 and 1000 m/z are collected, as this range will likely contain the putative small lipophilic ligands of interest. Inclusion of a small amount of water and/or acid can help in ionization, but may be omitted. Solvent only and resin blanks are run prior to any samples to aid in identification and elimination of common solvent impurities and resin contaminants.

If the activity is sufficiently pure to give a clear indication of a candidate ligand, the exact mass of the compound is determined (measurements made within 5 ppm) and a molecular formula is derived. Unless indicated otherwise by the distribution and intensities of isotopic peaks, we constrain molecular formula searches to the common organic elements C, H, O, N, S, and P. In the absence of acid, sodiated ions ($M+23$) are frequently the predominant species, so Na should also be included.

For more complex spectra containing multiple substances, or peaks of weak ion intensities with limited quantities of sample, we have successfully obtained simultaneous exact mass measurements of multiple peaks by calibrating the spectra with several internal standards. Caution should be applied in assignment of mass peaks to HPLC chromatographic data simply on the basis of comparing ion intensity with UV absorbance since in electrospray the ionization of a compound is determined by its functional groups and chemical nature. Therefore, the detection limits of different compounds vary substantially in a manner different from their optical properties. In addition, ionization of impurities may competitively inhibit the detection of sample components when injected in the direct ES-MS mode, thus making interpretation of absolute ion intensities between fractions only semi-quantitative. Nevertheless, a comparison of mass spectra from adjacent HPLC fractions can be useful in eliminating those peaks that do not follow the activity profile and in prioritizing candidate compounds. If available, an in-line narrow bore HPLC system coupled to ES-MS is ideal for maintaining sensitivity, subtraction of solvent, and contaminant background spectra and when attempting to match mass peaks to specific UV chromatographic peaks. We use a Micromass Q-TOF2 instrument with an Agilent 1100 series HPLC system for these purposes. This mass spectrometer has the added benefit of providing structural information when operated in the MS-MS mode that can be compared with putative synthetic standards of the candidate ligands.

Electron-Impact Mass Spectrometry

Electron-impact (EI) fragmentation spectra are highly reproducible and predictive with specific functional groups and substructures yielding distinctive fragmentation patterns. Therefore, the ability to generate such spectra on the target compounds can be highly beneficial. EI spectra can be screened against mass spectral databases, e.g., Wiley, or used to reconstruct a compound's substructures by looking for the sequential loss of probable fragmentation ions from the molecular ion. In addition, analysis by gas chromatography-mass spectrometry (GC-MS) allows collection of high-quality spectra of individual components from mixtures of compounds. Our primary instrument for these purposes is a Thermo-Finnegan Trace MS system. Samples are initially captured onto a DB-5 fused silica capillary column and components are sequentially volatized with a temperature gradient (linear from 50 to 290°C at 10°C/min, hold at 290°C for 20 min). Complex mixtures of closely related compounds can be resolved by retention time, and the method provides for an additional principle of separation distinct from solvent elution times from HPLC columns. Prior to

analysis, HPLC fractions must undergo solvent exchange into a nonpolar solvent, e.g., hexane or methylene chloride, to avoid undue damage to the GC column. Samples are evaporated under a stream of dry nitrogen or rotor evaporated, resuspended in 100–500 μl of dry solvent, and any residual water removed by drying the sample over anhydrous sodium sulfate. For very dilute samples, the solvent volume can be reduced further. Samples dissolved in as little as 10 μl can be used with conical autosampler vials. GC analysis is particularly well suited for the analysis of small volatile lipophilic molecules, but does have important limitations with respect to certain types of natural product characterizations. Thermal instability and excessive fragmentation may prevent collection of useful spectra. In addition, natural products containing multiple polar groups, e.g., those with more than one hydroxyl group, may be difficult to volatize. Carboxylic acids are refractory generally to analysis unless derivatized. Labile hydroxyl groups should be derivatized to form, for example, trimethylsilyl (TMS) ethers or other suitable groups to improve volatility, thermal stability, and fragmentation characteristics. Since derivatization is likely to destroy biological activity, it should be the last step in the analytical chain of purification. Numerous derivatization protocols can be found in any good analytical text.[13]

Concluding Remarks

The recent identification of orphan receptor ligands has increased greatly our understanding of such important processes as cholesterol and bile acid metabolism and the xenobiotic response. Most of the remaining orphan receptors have apparent ligand-binding pockets, hence it is likely that there are several ligand-binding pockets with endogenous ligands. If past trends continue, they will lead to surprising new insights into the homeostatic regulation of cellular biochemistry and novel signaling pathways. Perhaps, a perceived lack of expertise underlies the reluctance of molecular, cellular, and developmental biologists to venture into unfamiliar scientific disciplines, such as natural product isolation and structure determination. However, as access to easy-to-use advanced analytical instrumentation (e.g., mass spectrometry) becomes more commonplace, so too should the desire to incorporate these methods into the standard repertoire of techniques available within any modern biology laboratory. We hope that the strategies outlined above serve to illustrate that ligand identification using activity-guided purification is within the reach of most laboratories interested in fully characterizing nuclear hormone receptor biology.

[13] D. R. Knapp, "Handbook of Analytical Derivatization Reactions," 1st Ed. Wiley-Interscience, John Wiley & Sons, Inc., New York, 1979.

Acknowledgments

We thank Drs. Jochen Buck, Fadila Derguini, William Fenical, Heonjoong Kang, and John Greaves for past advice on natural product extraction, purification, and characterization. Research on receptor ligand identification in the author's laboratory was supported by grants from the US EPA (STAR G9D1 0090) and the NIH (GM60572).

[2] Quantitation of Receptor Ligands by Mass Spectrometry

By ERIK G. LUND and ULF DICZFALUSY

Biological Significance and Available Methods

The human genome contains approximately 50 members of the nuclear receptor family.[1] It is now clear that members of this class of proteins serve as important signal transducers, responding to the appearance of signaling molecules with an increase or decrease in the cellular expression of target genes. This behavior is typified by the classic steroid hormone receptors, which were among the first nuclear receptors to be identified and characterized.

Nuclear receptors can be subdivided into several classes, and according to one subdivision they may be classified as either being endocrine receptors, adopted orphan receptors, or orphan receptors.[2] The first group consists of well-characterized endocrine hormone receptors such as the vitamin A and D receptors, and the steroid and thyroid hormone receptors. These proteins typically have very high-affinity specific ligands. The second group consists of receptors for dietary or endogenous ligands present at higher concentration, which often bind several members of a class of molecules as ligands. This group is exemplified by the PPAR$\alpha/\delta/\gamma$ receptors, which bind fatty acids and their derivatives, the LXRα/β receptors, which bind oxysterols, and FXR, which interacts with bile acids. The third group consists of receptors without identified ligands, and it is possible that many receptors of this class actually do not have endogenous ligands. With the identification of ligands for many different nuclear receptors, a pattern can be recognized in that they are all generally relatively small, lipophilic compounds, reflecting the necessity for extracellular ligands to cross the plasma membrane for access to the receptor. This characteristic distinguishes nuclear receptors from cell surface receptors, which often have proteins or peptides as natural ligands.

[1] J. M. Maglich, A. Sluder, X. Guan, Y. Shi, D. D. McKee, K. Carrick, K. Kamdar, T. M. Willson, and J. T. Moore, *Genome Biol.* **2**, 1 (2001).
[2] A. Chawla, J. J. Repa, R. M. Evans, and D. J. Mangelsdorf, *Science* **294**, 1866 (2001).