Jennifer K. Colucci and Eric A. Ortlund*Department of Biochemistry and Winship
Cancer Institute, Emory University School of
Medicine, Atlanta, GA 30033, USA

Correspondence e-mail: eortlun@emory.edu

Received 1 May 2013

Accepted 5 July 2013

Expression, purification and crystallization of the ancestral androgen receptor–DHT complex

Steroid receptors (SRs) are a closely related family of ligand-dependent nuclear receptors that mediate the transcription of genes critical for development, reproduction and immunity. SR dysregulation has been implicated in cancer, inflammatory diseases and metabolic disorders. SRs bind their cognate hormone ligand with exquisite specificity, offering a unique system to study the evolution of molecular recognition. The SR family evolved from an estrogen-sensitive ancestor and diverged to become sensitive to progestagens, corticoids and, most recently, androgens. To understand the structural mechanisms driving the evolution of androgen responsiveness, the ancestral androgen receptor (ancAR1) was crystallized in complex with 5 α -dihydrotestosterone (DHT) and a fragment of the transcriptional mediator/intermediary factor 2 (Tif2). Crystals diffracted to 2.1 Å resolution and the resulting structure will permit a direct comparison with its progestagen-sensitive ancestor, ancestral steroid receptor 2 (AncSR2).

1. Introduction

The androgen receptor (AR) is a member of the steroid receptor (SR) family of transcription factors, which play a major role in regulating vertebrate biology. AR responds to androgens, such as testosterone and 5 α -dihydrotestosterone (DHT), to regulate genes central to male sexual development, immunity and behavior (Li *et al.*, 2012; Holterhus, 2011; Walters *et al.*, 2010; Wang *et al.*, 2009). Given the widely prevalent role of androgens in normal physiology, AR signaling has been implicated in a number of diseases including cancer, cardiovascular defects, metabolic disorders, Alzheimer's and androgen insensitivity syndrome (AIS) (Li *et al.*, 2012; Hickey *et al.*, 2012; Hughes *et al.*, 2012). AR plays a particularly malicious role in prostate cancer by driving gene expression to fuel cell growth in both an androgen-dependent and an androgen-independent manner (Green *et al.*, 2012; Tamburrino *et al.*, 2012).

AR displays the typical modular SR domain architecture with an N-terminal activation function 1 domain, a DNA-binding domain, a short linker region and a ligand-binding domain (LBD). Without ligand, AR is unstable and resides in the cytoplasm complexed to chaperones (Fang *et al.*, 1996). Upon binding to a high-affinity ligand such as DHT, the hormone–receptor complex translocates to the nucleus, where it binds to coregulatory proteins such as Tif2 (transcriptional mediator/intermediary factor 2) to regulate target gene expression (Nagy & Schwabe, 2004). This simple paradigm, in which a small lipophilic ligand regulates complex gene programs, requires AR to recognize androgens with a high degree of specificity and exclude interaction with very similar steroids such as estrogens, progestagens and corticosteroids. This exquisite sensitivity to androgens arose during early vertebrate evolution with the appearance of the first AR (some 450 million years ago) following the duplication and subsequent divergence from a progestagen-activated ancestor (Eick & Thornton, 2011; Ogino *et al.*, 2009).

To gain insight into the fundamental processes governing the evolution of androgen specificity, we initiated biophysical studies on the ancestral AR, ancAR1, the sequence of which was inferred using a well described technique termed ancestral gene resurrection (to be published; Thornton, 2004; Harms & Thornton, 2010; Ortlund *et al.*,

2007). We have cloned, expressed, purified and crystallized ancAR1 in complex with the high-affinity ligand DHT and a fragment of the transcriptional coactivator Tif2 which binds to the activated conformation of SRs. Diffraction data were collected to 2.1 Å resolution.

2. Materials and methods

2.1. Reagents

Chemicals were purchased from Sigma (St Louis, Missouri, USA) or Fisher (Hampton, New Hampshire, USA). DHT was purchased from Toronto Research Chemicals (Toronto, Canada). The vector for His-tagged TEV protease was a gift from David Waugh (National Cancer Institute). The pLIC_MBP vector was a gift from John Sondek (UNC, Chapel Hill). The ancAR1 LBD was resurrected using well established protocols and was kindly provided by Dr Joseph Thornton (University of Oregon, USA). The peptide corresponding to the nuclear receptor coactivator box 3 from human Tif2 was synthesized by RS Synthesis (Louisville, Kentucky, USA).

2.2. Cloning

The ancAR1 LBD (residues 1–250) was cloned into pLIC_MBP, which contains a hexahistidine tag followed by the maltose-binding protein (MBP) and a *Tobacco etch virus* (TEV) protease site N-terminal to the protein. The forward cloning primer used was 5'-TACTTCCAATCCAATGCGGCGATCGCCATTCCCATTTTCC-3'; the reverse cloning primer used was 5'-TTATCCACTTCCAATGCGCTAGTTTAACTTACTGC-3'. The sequence of ancAR1 is IPIFLSVLSQIEPEVVYAGYDNTQPDTSASLLTSLNELGERQLVRVVKWAKALPGFRNLHVDDQMTLIQYSWMGVMVFAMGWRSYKNVNSRMLYFAPDLVFNEORMQKSTMYNLCVVRMRHLSQEFVWLQVTOEEFLCMKALLFSIIPVEGLKNQKYFDELRMNYIKELDRVISFOGKNPTSSQRFYQLTKLLDSLQPIVRKLHQFTFDLFVQSQSLSVEFPPEMMSEIISAQVPKILAGMVKPLLFHKQ. The crystallization construct contains the residues SNA C-terminal to the receptor as a relic from the TEV protease cleavage site.

2.3. Expression and purification

AncAR1 LBD was expressed as a 6×His-MBP fusion protein in *Escherichia coli* BL21(DE3) cells. Cultures (1.0 l in Terrific Broth)

Figure 1
Following a series of affinity columns, ancAR1–DHT was purified to homogeneity. Lane 1, TEV protease-cleaved ancAR1-MBP fusion protein. Lane 2, purified ancAR1. Lane 3, cleaved MBP. Lane M contains molecular-weight markers (labeled in kDa).

Table 1

Diffraction and processing statistics for the ancAR1–DHT crystal.

Values in parentheses are for the outermost shell.

Resolution (Å)	2.10 (29.40–2.10)
Space group	$P4_32_12$
Unit-cell parameters (Å, °)	$a = b = 68.9$, $c = 147.3$, $\alpha = \beta = \gamma = 90$
No. of reflections	172781
$R_{\text{merge}}^{\dagger}$ (%)	10.9 (43.7)
Completeness (%)	99.9 (100)
Average multiplicity	8.0 (7.3)
$\langle I/\sigma(I) \rangle$	17.4 (4.3)
Mosaicity (°)	0.45

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of several symmetry-related observations.

were grown to an OD₆₀₀ of 0.8 and induced with a final concentration of 400 μM IPTG and 50 μM DHT at 291 K overnight. Cell mass was collected by centrifugation at 4000 rev min^{−1} for 20 min, resuspended in 150 mM NaCl, 20 mM Tris–HCl pH 7.4, 5% glycerol, 25 mM imidazole, 0.1% PMSF and lysed using sonication on ice. AncAR1-MBP was initially purified using Ni²⁺-affinity chromatography (HisTrap column, GE Healthcare). Fractions containing ancAR1-MBP were identified by denaturing polyacrylamide gel electrophoresis (SDS–PAGE), pooled and dialyzed against 150 mM NaCl, 20 mM Tris–HCl pH 7.4, 5% glycerol, 1 mg TEV protease. Following TEV protease cleavage, the tagged MBP was removed by an additional Ni²⁺-affinity column. The flowthrough containing ancAR1 LBD was concentrated using an Amicon Ultra 10K centrifugal filter device (Millipore), concentrated to 3.3 mg ml^{−1} and dialyzed against 150 mM sodium chloride, 20 mM Tris–HCl pH 7.4, 5% glycerol. The final purity of the ancAR1 LBD was assessed using SDS–PAGE (Fig. 1).

2.4. Crystallization and data collection

Prior to crystallization, an additional 50 μM DHT was added to the ancAR1–DHT complex to ensure full occupancy of DHT in the ligand-binding pocket. Additionally, 500 μM of a peptide derived from human Tif2, corresponding to the nuclear receptor coactivator box 3 (740-KENALLRYLLDKDD-753), was added to the receptor–ligand complex for crystallization, yielding a 5:1 molar ratio of peptide:ancAR1. Crystallization trials were performed using sitting-drop vapor diffusion, mixing 0.2 μl of the protein sample with an equal volume of screening solution and equilibrating against 60 μl screening solution in the reservoir. Initial screening was performed using 480 conditions from the commercially available kits The JCSG+ Suite, The PEGs Suite, The Nucleix Suite, The Classics Lite Suite and The AmSO₄ Suite (Qiagen). Positive hits were obtained using The PEGs Suite condition A11 [0.1 M MES pH 6.5, 25% (w/v) PEG 1000], The JCSG+ Suite condition A2 (0.1 M trisodium citrate pH 5.5, 20% PEG 3000) and The JCSG+ Suite condition G9 (0.1 M KCN, 30% PEG MME 2000). These hits were further expanded to generate diffraction-quality crystals.

Crystals of the ternary ancAR1 LBD–DHT–Tif2 complex were grown by hanging-drop vapor diffusion at 295 K from solutions consisting of 1.0 μl protein at 3.3 mg ml^{−1} and 1.0 μl of the crystallant 20% PEG 1000, 0.3 M MES pH 6.5. Crystals were cryoprotected by transient soaking in crystallant containing 20% glycerol and were flash-cooled in liquid N₂ at 100 K. 205 frames were collected at 0.5° oscillation. Data to 2.1 Å resolution were collected on the South East Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source (APS) at Argonne National Laboratory in Chicago, Illinois, USA using a wavelength of 0.97 Å and were

Figure 2
Crystals of ancAR1–DHT. The long rod-shaped crystals are approximately 100–200 μm in length. The crystals grew in 20% PEG 1000, 0.3 M MES pH 6.5.

Figure 3
Diffraction image of an ancAR1–DHT crystal. The detector edge corresponds to 2.15 Å resolution.

processed and scaled with *HKL-2000* (Otwinowski & Minor, 1997) (Table 1). Data processing revealed that the crystals of the ancAR1–DHT complex grew in space group $P4_32_12$.

3. Results and discussion

In order to understand the evolution of ligand specificity in steroid hormone nuclear receptors, we cloned, overexpressed and purified ancAR1 in complex with its most potent ligand, DHT, and a fragment of the human coactivator Tif2. A denaturing SDS–PAGE gel shows a pure receptor–ligand complex with no contaminating bands (Fig. 1). The calculated molecular weight of the protein is 29 498 Da.

Crystals were grown by equilibrating 1.0 μl protein solution and 1.0 μl mother liquor using hanging-drop vapor diffusion (Fig. 2). The receptor crystallized as long rods in a solution consisting of 20% PEG 1000, 0.3 M MES pH 6.5. Crystals formed in space group $P4_32_12$ and diffracted to 2.1 Å resolution (Fig. 3). The Matthews coefficient (V_M) was 2.96 Å³ Da^{−1} with one monomer in the asymmetric unit, corresponding to a solvent content of 58.4% (Matthews, 1968).

Solving the crystal structure of this ancient receptor–ligand complex is a critical step towards understanding the structural and biophysical changes that occurred in the AR lineage to develop sensitivity to androgenic compounds. This structure will permit direct structural comparison with the progestagen-activated ancestral steroid receptor 2 (Eick *et al.*, 2012) and will guide future functional studies identifying the residues that were responsible for the functional shift from 17-acetyl steroid to 17-hydroxyl (androgen) responsiveness. Understanding how ligand recognition can be evolved, harvested and exploited is essential to the progression of protein engineering, drug design and discovery, and a true comprehension of our molecular history.

We would like to thank Dr Geeta Eick and Dr Joseph W. Thornton (University of Oregon, USA) for providing the ancAR1 gene. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at <http://www.ser-cat.org/members.html>. This research is supported by R01 GM081592 (Joseph Thornton PI, EAO coPI), start-up funds from Emory University (EAO) and JKC's American Heart Association pre-doctoral fellowship (10PRE3530007). Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38.

References

- Eick, G. N., Colucci, J. K., Harms, M. J., Ortlund, E. A. & Thornton, J. W. (2012). *PLoS Genet.* **8**, e1003072.
- Eick, G. N. & Thornton, J. W. (2011). *Mol. Cell. Endocrinol.* **334**, 31–38.
- Fang, Y., Fliss, A. E., Robins, D. M. & Caplan, A. J. (1996). *J. Biol. Chem.* **271**, 28697–28702.
- Green, S. M., Mostaghel, E. A. & Nelson, P. S. (2012). *Mol. Cell. Endocrinol.* **360**, 3–13.
- Harms, M. J. & Thornton, J. W. (2010). *Curr. Opin. Struct. Biol.* **20**, 360–366.
- Hickey, T. E., Robinson, J. L., Carroll, J. S. & Tilley, W. D. (2012). *Mol. Endocrinol.* **26**, 1252–1267.
- Holterhus, P. M. (2011). *Pediatr. Endocrinol. Rev.* **9**, Suppl. 1, 515–518.
- Hughes, I. A., Werner, R., Bunch, T. & Hiort, O. (2012). *Semin. Reprod. Med.* **30**, 432–442.
- Li, Y., Izumi, K. & Miyamoto, H. (2012). *Jpn. J. Clin. Oncol.* **42**, 569–577.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nagy, L. & Schwabe, J. W. (2004). *Trends Biochem. Sci.* **29**, 317–324.
- Ogino, Y., Katoh, H., Kuraku, S. & Yamada, G. (2009). *Endocrinology*, **150**, 5415–5427.
- Ortlund, E. A., Bridgham, J. T., Redinbo, M. R. & Thornton, J. W. (2007). *Science*, **317**, 1544–1548.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Tamburrino, L., Salvianti, F., Marchiani, S., Pinzani, P., Nesi, G., Serni, S., Forti, G. & Baldi, E. (2012). *Steroids*, **77**, 996–1001.
- Thornton, J. W. (2004). *Nature Rev. Genet.* **5**, 366–375.
- Walters, K. A., Simanainen, U. & Handelsman, D. J. (2010). *Hum. Reprod. Update*, **16**, 543–558.
- Wang, R.-S., Yeh, S., Tzeng, C.-R. & Chang, C. (2009). *Endocr. Rev.* **30**, 119–132.