

# Directed Evolution of Orthogonal Ligand Specificity in a Single Scaffold\*\*

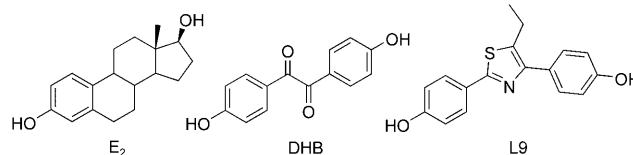
Michael J. McLachlan, Karuppiah Chockalingam, Ka Chun Lai, and Huimin Zhao\*

Gene-regulation systems that provide temporal and spatial regulation of target-gene expression in response to small-molecule ligands (small-molecule-dependent gene switches or circuits) are powerful tools for gene therapy, tissue engineering, metabolic engineering, and functional genomics.<sup>[1,2]</sup> Notably, small-molecule-dependent gene switches were recently used to generate induced pluripotent stem cells (iPS) for regenerative medicine.<sup>[3,4]</sup> The need for orthogonal regulatory elements is evidenced by the surge in interest in the creation of gene circuits, inspired by their electrical counterparts.<sup>[5]</sup> Numerous orthogonal small-molecule-dependent gene switches have been developed to exert control over transcription, translation, or protein function.<sup>[6]</sup> However, most either use bacterial components that involve regulation of a specific operator sequence bound by an antibiotic-regulated repressor, or use nuclear-receptor mutants that respond to synthetic hormones, such as RU486, which in large doses affect human physiology. An ideal gene-switch system would be nonimmunogenic, would show promoter flexibility through different DNA-binding domains, and would be expandable to new inducers.

Nuclear hormone receptors (NHRs) offer desirable characteristics as gene switches for transcriptional control. The binding of a small ligand results in NHR dimerization, translocation, and the activation of promoters that harbor specific responsive elements.<sup>[7]</sup> With distinct domains for ligand binding, DNA binding, and activation/repression functions, NHRs offer protein engineers the flexibility to create chimeric transcriptional activators or repressors by modular design.<sup>[8,9]</sup> The ability to access small-molecule hormonelike compounds<sup>[10]</sup> by organic synthesis makes these natural allosteric transcriptional switches attractive targets for engineering. Nuclear hormone receptors enable a wide range of target-protein-expression levels, which are tunable through variation of the dose of a ligand. NHR

ligand-binding domains (LBDs) have also been combined with other proteins to enable posttranslational control of protein function.<sup>[11]</sup> Specificity-reengineering approaches involving NHRs have typically involved mutation of the LBD to a form that is not activated by the natural ligand but instead by a synthetic small molecule that is inactive against the wild-type LBD. The LBD of the human estrogen receptor  $\alpha$  (hER $\alpha$ ) has proven to be a particularly versatile platform for the creation of orthogonal ligand–receptor pairs.<sup>[12,13]</sup> Despite these efforts and studies on other NHR LBDs, the creation of unique, NHR-based independently functioning ligand–receptor pairs that are not only orthogonal to cellular elements but that do not cross-interact with one another has yet to be demonstrated.

Previously,<sup>[14]</sup> we identified residues in the ligand-binding pocket of the hER $\alpha$  LBD that are important for ligand specificity. After four cycles of saturation mutagenesis and one cycle of random mutagenesis we identified two hER $\alpha$  LBD variants, 4S and 5E, which are activated by 4,4'-dihydroxybenzil (DHB) but not 17 $\beta$ -estradiol (E<sub>2</sub>; Scheme 1). To further increase the sensitivity of mutant 4S



Scheme 1. Chemical structures of E<sub>2</sub>, DHB, and L9.

to DHB, 11 additional sites were chosen for further mutagenesis: 349, 387, 391, 404, and 524 from within the ligand-binding pocket, and 442, 459, 466, 534, 536, 537 from outside the ligand-binding pocket. Three rounds of stepwise site-saturation mutagenesis on these plus the remaining 343, 347, 383, 384, 421, 424, 425, 428, 525, and 528 LBD sites were carried out on 4S to give the final mutant 7S (see Figure S1 and Table S1 in the Supporting Information).

To demonstrate that our strategy for generating sensitive and specific ligand–receptor pairs is generally applicable, we created mutants responsive toward 2,4-di(4-hydroxyphenyl)-5-ethylthiazole (L9; Scheme 1). Six rounds of stepwise site-saturation mutagenesis were applied to 19 residues of the wild-type hER $\alpha$  LBD (WT): 343, 346, 347, 349, 350, 383, 384, 387, 388, 391, 404, 421, 424, 425, 428, 521, 524, 525, and 528. Only two of the 21 ligand-binding-pocket residues,<sup>[14]</sup> 353 and 394, were left out, owing to their critical role in maintaining the orientation of the A ring of the ligand. As in the approach used for the creation of DHB-specific mutants, parallel

[\*] M. J. McLachlan,<sup>[†]</sup> K. Chockalingam,<sup>[†]</sup> K. C. Lai, Prof. H. Zhao  
Department of Chemical and Biomolecular Engineering  
University of Illinois at Urbana-Champaign  
600 S. Mathews Avenue, Urbana, IL 61801 (USA)  
E-mail: zhao5@illinois.edu  
Homepage: <http://www.chemeng.uiuc.edu/~zhaogr>

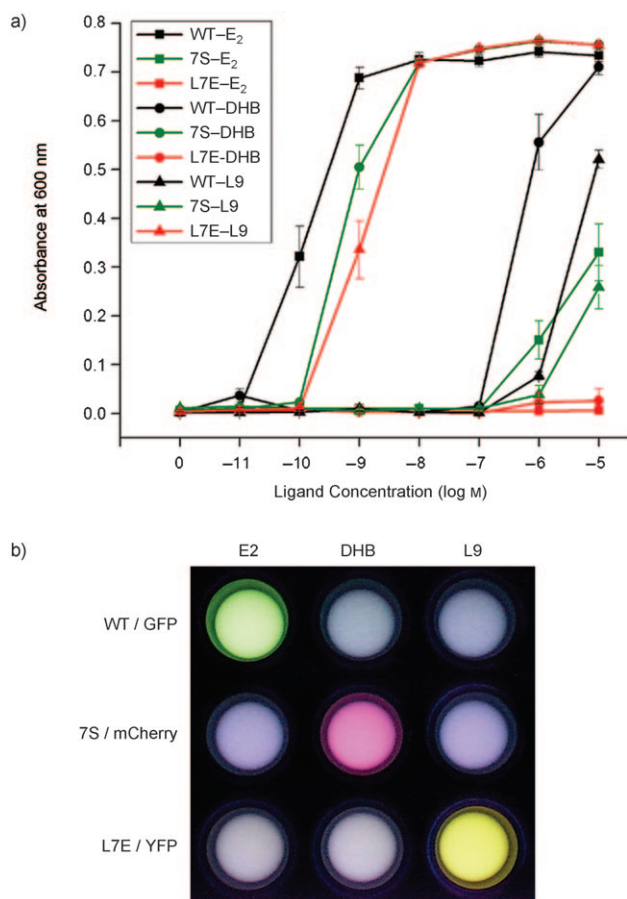
[†] These authors contributed equally.

[\*\*] We thank Ee-Lui Ang, Meng Chen, and Victor Gonzalez for technical assistance and helpful discussions. We also thank Carlos Barbas III for the gift of a plasmid containing the KRAB repressor, and Yingxiao Wang and Roger Tsien for the gift of mCherry and YFP plasmids. This research was supported by a National Science Foundation CAREER Award (BES-0348107).

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.200903413>.

positive screening for L9 response and negative screening against  $E_2$  was performed in each round as described previously.<sup>[14]</sup> Site-saturation mutagenesis was followed by one round of error-prone PCR accompanied by growth-based selection for further L9-specific mutants to give the final mutant L7E. We determined the sensitivity of the best mutant after each round toward the ligands L9,  $E_2$ , and DHB (see Figure S1 and Table S1 in the Supporting Information).

Remarkable sensitivity and specificity were observed with both ligand–receptor pairs. Both showed nanomolar sensitivity toward their respective ligands but were not as sensitive as the native estrogen receptor toward its ligand  $E_2$  (Figure 1 a;

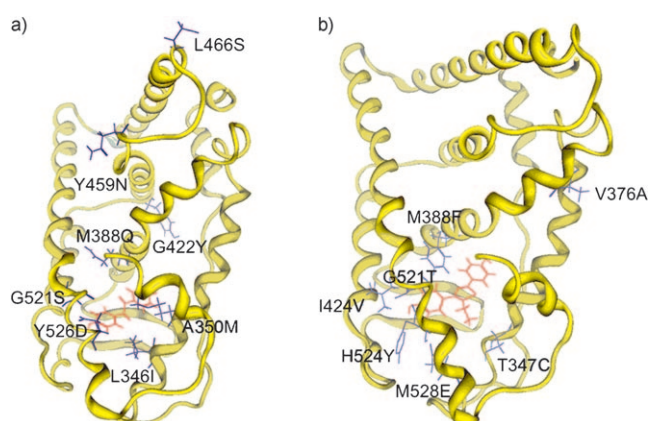


**Figure 1.** a) Dose-response curves for WT, 7S, and L7E in the presence of  $E_2$ , DHB, or L9. b) Control by gene switches of the expression of fluorescent proteins in yeast from a *Gal1* promoter. Samples in the top row contain GFP and WT, those in the middle row contain mCherry and 7S, and those in the bottom row contain YFP and L7E. The three columns contain  $E_2$ , DHB, or L9 (from left to right) at a concentration of  $10^{-7}$  M.

see also Table S1 in the Supporting Information). The mutant 7S was 850 times more sensitive than the WT toward DHB, and around 300 000 times less sensitive toward  $E_2$ . Thus, an approximately  $2.5 \times 10^9$ -fold specificity shift was observed. This mutant only showed a response to L9 at micromolar concentrations. Although 7S exhibited a stronger response to  $E_2$  than the parent mutant 4S, the overall selectivity for DHB

over  $E_2$  was maintained or slightly improved upon through the three rounds of mutagenesis. The mutant L7E was 3400 times more sensitive to L9 than the WT and showed no response to  $E_2$  or DHB at concentrations up to  $10^{-5}$  M. Although the receptors share the same protein scaffold, negligible cross-reactivity was observed between the WT– $E_2$ , 7S–DHB, and L7E–L9 receptor–ligand pairs. The wild-type, 7S, and L7E LBDs were activated specifically by their respective ligands when used to control the expression of green fluorescent protein (GFP), mCherry, and yellow fluorescent protein (YFP) in yeast (Figure 1 b).

The mutations in 7S and L7E were mapped onto a crystal structure of the wild-type hER $\alpha$  LBD (Figure 2). The new mutations that are not present in 4S are situated outside the

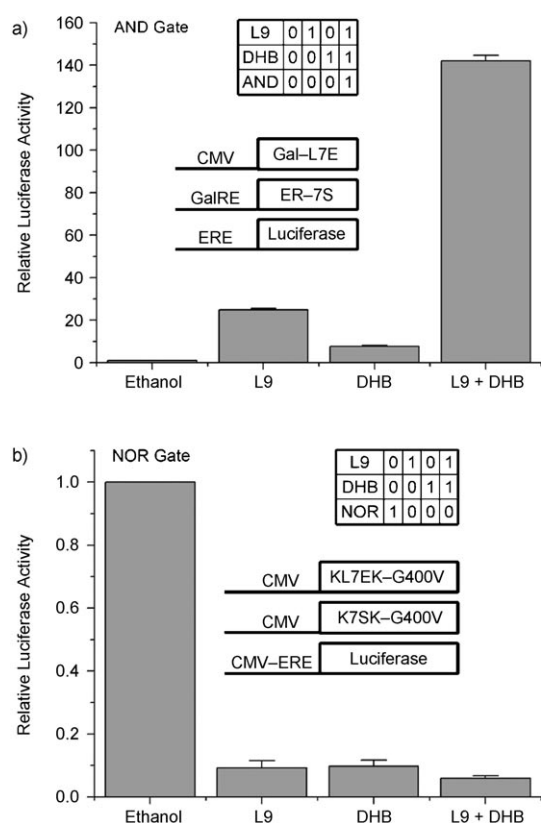


**Figure 2.** The locations of mutated sites in a) mutant 7S and b) mutant L7E. The DHB and L9 ligands were docked into the cavity left by  $E_2$  in the crystal structure of hER $\alpha$  LBD– $E_2$  (PDB code: 1GWR), as described elsewhere.<sup>[14]</sup>

ligand-binding pocket. Position 442 is located at the N terminus of helix 8. It has been suggested that this residue indirectly influences the ligand-binding and transactivation activity of the estrogen receptor.<sup>[15]</sup> Residue 459 is situated in the dimerization interface of the homodimer; thus, Y459N may enhance the ability of the mutant to dimerize efficiently in the presence of DHB. The role of residue 466 is not immediately clear. A pivotal role was observed for residue 521 in the determination of ligand specificity: the first mutation found for the L9 ligand, G521T, caused an almost  $10^4$ -fold shift in specificity toward L9 over  $E_2$ . Interestingly, G521S was responsible for a significantly shifted specificity toward DHB for the DHB-specific mutant 4S.<sup>[14]</sup> Another mutation at this site, G521R, is known to reduce the affinity of the hER $\alpha$  LBD toward  $E_2$  by a factor of more than  $10^4$ .<sup>[16]</sup> A frame-shifting deletion at position 577 in the mutant L7E was removed in further cloning with no detectable change in ligand responsiveness (data not shown). The error-prone PCR mutation V376A in L7E may confer an improved ability to bind coactivators. The 376 site has been implicated in the formation of a groove with helix 12 of the hER $\alpha$  LBD; this groove enables the docking of coactivator proteins during transcription.<sup>[17]</sup> Molecular-dynamics simulations may pro-

vide clues as to the effect of the mutations; however, such simulations are most effective when a structure is available.

To demonstrate the utility of our engineered ligand–receptor pairs for the creation of genetic circuits in eukaryotic systems, we implemented the logic functions AND and NOR in HeLa cells by using a luciferase reporter, Gal4- or estrogen-receptor DNA-binding-domain promoter-response elements (GalRE or ERE), wild-type or engineered LBDs, and the KRAB (Krüppel-associated box) repressor domain. The AND gate (Figure 3a) was constructed with a constitu-



**Figure 3.** a) Logic gates in HeLa cells. The graphs show luciferase activity (mean, standard error) normalized to  $\beta$ -galactosidase expression, and with the value for treatment with ethanol taken as equal to 1. a) AND gate with L9 ( $10^{-8}$  M), DHB ( $10^{-8}$  M), or both (each  $10^{-8}$  M). b) NOR gate with L9 ( $10^{-6}$  M), DHB ( $10^{-6}$  M), or both (each  $10^{-6}$  M).

tively expressed Gal4L7E switch, a GalRE-controlled ER-7S switch, and an ERE-controlled luciferase gene. Induction of luciferase was mainly observed when both L9 and DHB ligands were present (140-fold), although some leaky expression was observed when either ligand was present alone. The NOR gate (Figure 3b) was comprised of constitutively expressed KRAB-ER-L7E-KRAB-G400V or KRAB-ER-7S-KRAB-G400V switches, with luciferase downstream of a constitutive cytomegalovirus (CMV) promoter with an inserted ERE promoter-response element. This circuit showed 10–16-fold repression in the presence of L9, DHB, or both ligands.

In summary, a single protein scaffold was used in an effective and repeatable manner for the generation of two highly sensitive orthogonal ligand–receptor pairs. A more than  $2 \times 10^9$ -fold specificity shift was observed for both ligand–receptor pairs, each of which is unique in its ligand specificity. Despite a scarcity of examples of receptor-based logic gates,<sup>[5,18]</sup> our engineered ligand–receptor pairs were combined successfully into higher-order functions, as exemplified by the creation of logic gates in mammalian cells. It is envisioned that any number of further pairs could be generated in the same way for the development of new small-molecule-regulated gene-expression systems. These results add to the field of specificity reengineering, which has been applied to other protein scaffolds, such as lipocalins and antibodies.<sup>[20]</sup>

## Experimental Section

**Yeast two-hybrid screening:** The protocols used were described previously.<sup>[14]</sup> For further engineering of 4S, the choice of the six residues outside the ligand-binding pocket was based on their observed role in increasing the potency of response toward different synthetic ligands in the context of different LBD-mutant templates. Of these six non-ligand-binding-pocket sites, 442 and 536 were chosen on the basis of previous observations of their importance in the promiscuous increase of the transactivation response of mutant estrogen receptors toward different ligands,<sup>[15,19]</sup> 537 was chosen on the basis of its role in affecting the strength of ligand response, as determined by studies in our laboratory (data not shown), and the remaining three sites were chosen on the basis of site-directed-mutagenesis studies.

**Fluorescent yeasts:** Plasmids (gal-GFP + pBD-Gal4-WT + pGAD424-SRC1; gal-mCherry + pBD-Gal4-7S + pGAD424-SRC1; gal-YFP + pBD-Gal4-L7E + pGAD424-SRC1) were transformed into the yeast strain YM4271 (Clontech, Mountain View, CA) and grown on SC-LWU (synthetic complete media minus leucine, tryptophan, and uracil) plates. Colonies were grown overnight in SC-LWU medium at 30 °C, and an aliquot of the medium (200  $\mu$ L) was then inoculated into YPAD medium (2 mL) with the ligand ( $10^{-7}$  M) and grown overnight. Cells were pelleted, washed with phosphate-buffered saline (PBS), resuspended in PBS (50  $\mu$ L), and transferred to a 96-well plate. The plate was photographed on a UV illuminator (Spectroline, Westbury, NY) that produced light of wavelength 435 nm.

**Cell culture and transformation:** HeLa cells were grown in minimal essential medium (MEM) plus sodium pyruvate (1 mM) and 10 % fetal bovine serum (UIUC Cell Media Facility) at 37 °C with 5 % CO<sub>2</sub>. Cells were transferred to 24-well plates with media containing 5 % charcoal-dextran-stripped calf serum (UIUC Cell Media Facility), grown overnight, and transfected with lipofectamine 2000 (1.5  $\mu$ L; Invitrogen, Carlsbad CA),  $\beta$ -galactosidase expression plasmid (100 ng), luciferase reporter plasmid (690 ng), and the relevant gene-switch plasmid (10 ng). After 4 h, the medium was replaced, and the ligand E<sub>2</sub>, DHB, or L9 was added. Cells were incubated for 24 h, then lysed and assayed for luciferase activity with the Luciferase Assay System (Promega, Madison WI). Luciferase levels were normalized to  $\beta$ -galactosidase expression.

**Cloning of expression constructs:** The ERE-controlled luciferase reporter (2ERE-pS2-pGL3-Luc) has been used previously.<sup>[14]</sup> GalRE or the CMV promoter were added by blunt ligation upstream of the ERE element. Constitutively expressed gene switches were cloned between the KpnI and BamHI sites of pCMV5. Inducible gene switches were cloned by replacing the luciferase gene in the presence of the relevant promoter. The *GalI* promoter was amplified by PCR

from yeast genomic DNA and cloned upstream of GFP, mCherry, or YFP in pRS426.

Received: June 24, 2009

Published online: September 11, 2009

**Keywords:** directed evolution · gene technology · ligand specificity · molecular biology · proteins

- [1] M. Fussenegger, *Biotechnol. Prog.* **2001**, *17*, 1.
- [2] C. Toniatti, H. Bujard, R. Cortese, G. Ciliberto, *Gene Ther.* **2004**, *11*, 649.
- [3] M. Wernig, C. J. Lengner, J. Hanna, M. A. Lodato, E. Steine, R. Foreman, J. Staerk, S. Markoulaki, R. Jaenisch, *Nat. Biotechnol.* **2008**, *26*, 916.
- [4] S. Markoulaki, J. Hanna, C. Beard, B. W. Carey, A. W. Cheng, C. J. Lengner, J. A. Dausman, D. Fu, Q. Gao, S. Wu, J. P. Cassady, R. Jaenisch, *Nat. Biotechnol.* **2009**, *27*, 169.
- [5] D. J. Sayut, P. K. Kambam, L. Sun, *Mol. Biosyst.* **2007**, *3*, 835.
- [6] W. Weber, M. Fussenegger, *Handb. Exp. Pharmacol.* **2007**, *178*, 73.
- [7] A. Aranda, A. Pascual, *Physiol. Rev.* **2001**, *81*, 1269.
- [8] R. R. Beerli, U. Schopfer, B. Dreier, C. F. Barbas III, *J. Biol. Chem.* **2000**, *275*, 32617.
- [9] S. Braselmann, P. Graninger, M. Busslinger, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1657.
- [10] B. E. Fink, D. S. Mortensen, S. R. Stauffer, Z. D. Aron, J. A. Katzenellenbogen, *Chem. Biol.* **1999**, *6*, 205.
- [11] D. Picard, *Methods Enzymol.* **2000**, *327*, 385.
- [12] P. Gallinari, A. Lahm, U. Koch, C. Paolini, M. C. Nardi, G. Roscilli, O. Kinzel, D. Fattori, E. Muraglia, C. Toniatti, R. Cortese, R. De Francesco, G. Ciliberto, *Chem. Biol.* **2005**, *12*, 883.
- [13] J. Whelan, N. Miller, *J. Steroid Biochem.* **1996**, *58*, 3.
- [14] K. Chockalingam, Z. Chen, J. A. Katzenellenbogen, H. Zhao, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5691.
- [15] F. C. Eng, H. S. Lee, J. Ferrara, T. M. Willson, J. H. White, *Mol. Cell. Biol.* **1997**, *17*, 4644.
- [16] R. Paulmurugan, A. Tamrazi, J. A. Katzenellenbogen, B. S. Katzenellenbogen, S. S. Gambhir, *Mol. Endocrinol.* **2008**, *22*, 232.
- [17] A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, *95*, 927.
- [18] D. Greber, M. Fussenegger, *J. Biotechnol.* **2007**, *130*, 329.
- [19] Z. Chen, B. S. Katzenellenbogen, J. A. Katzenellenbogen, H. Zhao, *J. Biol. Chem.* **2004**, *279*, 33855.
- [20] C. Grönwall, S. Ståhl, *J. Biotechnol.* **2009**, *140*, 254.