PNAS
www.pnas.org

Supplementary Information for

Molecular Evolution of the Switch for Progesterone from Mineralocorticoid Receptor Agonist to Antagonist

Peter J. Fuller[a,], Yi-Zhou Yao[a], Ruitao Jin[b], Sitong He[b], Beatriz Martín-Fernández[a,2], Morag J. Young[a] and Brian J. Smith[b].

[a]Centre for Endocrinology and Metabolism, Hudson Institute of Medical Research and the Monash University Department of Molecular Translational Science, Clayton, Victoria 3168, Australia. [b]La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia.

Corresponding author: Peter J. Fuller
Email: peter.fuller@hudson.org.au

This PDF file includes:

Supplementary text
Figures S1 to S10
SI References

Supplementary Information Text

Materials and Methods

Construction of Chimeras. The previously described hMR expression vector pRShMR (**6**) was used as the starting point with the LBD being replaced by a series of hMR:zMR chimeras created between amino acids 672-984 of the hMR LBD and the corresponding amino acids 670-970 of the zMR LBD (*SI Appendix,* Fig. S1). These chimeras were named by giving each LBD chimera a four-letter name based on the sequence in each of four sections, where H is hMR sequence and Z is zMR sequence. Amino acids 785-880 of the hMR (zMR 771-866) correspond to the second of these four sections. Sub-chimeras were created within this second region in which hMR amino acids 785-843 were exchanged with zMR amino acids 771-829 in the zMR LBD and amino acids 844-880 from the hMR were exchanged with amino acids 830-866 in the zMR. Two reciprocal chimeras were also created in the hMR LBD. All chimeras were created by overlap extension PCR using pRShMR and pcDNA3.1zMR as templates. The PCR was performed using GoTaq Green Master Mix (Promega, Madison, CA). The chimeric sequence was subsequently removed

from the vector by digestion with XhoI and HindIII and ligated into pRShMR, digested with XhoI and HindIII, and then fully sequenced. The primers used for overlap extension PCR are shown in the supplementary data (*SI Appendix,* Fig. S1).

Construction of Point Mutations. Single, double, and multiple point mutations were made to incorporate combinations of amino acids from zMR (771-866) into the ZHZZ chimera or from hMR (785-880) into the HZHH chimera (*SI Appendix,* Fig. S1). These mutations were created using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA), as per the manufacturer's instructions. Successful mutagenesis was confirmed by direct sequencing. The primers used for the PCR mutagenesis are presented in (*SI Appendix,* Table S1).

Transactivation Assays in CV-1 cells. The MMTV-LUC reporter and Ren-LUC control plasmids have been described previously (**6**). CV-1 African green monkey cells were grown at 37^{o}C in DMEM supplemented with 0.075% sodium bicarbonate, 10 mM HEPES, 1 mM glutamine, non-essential amino acids and penicillin (10U/L)-streptomycin (10μg/L)-fungazone (0.025μg/L) + 10% foetal bovine serum (FBS). The cells were trypsinized and replated in 24-well plates at a density of 5 x 10^{4} cells/well in DMEM + 5% FBS. After 24 hours, the cells were transfected with 0.25μg of pRShMR or pRSzMR or the relevant chimera/mutated MR, 0.25μg of MMTV-LUC reporter plasmid and 0.05μg of pRen-LUC control plasmid using FuGENE®HD Transfection Reagent (Promega, Madison, WI, USA) as per the manufacturer's protocol and the media was changed to DMEM supplemented with 5% charcoal stripped FBS. The cells were then incubated at 37^{o}C for 24 hours. Fresh DMEM + 5% charcoal-stripped FBS was added to the cells before addition of steroid. The cells were incubated with steroid for 24 hours at 37^{o}C, and then harvested for assay. Measurement of transactivation was performed using the "Dual-Luciferase® Reporter Assay System" kit (Promega, Madison, WI, USA). The cells were incubated with 100μL passive lysis buffer for 15 minutes at room temperature. A 20μL aliquot was removed to a 96-well plate for assay, and relative light units were measured for 1 second in a Perkin Elmer 2103 Envision Multilabile Reader (PerkinElmer Life and Analytical Sciences, CT, USA). *Renilla* luciferase activity was determined to control for transfection efficiency. All measurements were performed on three replicates in at least two separate experiments.

Transactivation assays in ZF-4 cells. The zebrafish embryonic fibroblast cell line, ZF-4 (**14**), was generously provide by Dr M-C Keightley (Monash University). They were cultured at 28°C in DMEM:F12(1:1) supplemented with 2.438g/L sodium bicarbonate, 1 mM glutamine, and penicillin (10U/L)-streptomycin (10μg/L) fungazone (0.025μg/L) + 10% foetal bovine serum (FBS). The cells were trypsinized and replated in 24-well plates at a density of 1 x 10^{4} cells/well in DMEM + 10% FBS. After 24 hours the cells were transfected with 0.25μg of the MR expression plasmid, 0.25μg of MMTV-LUC reporter plasmid and 0.05μg of pRen-LUC control plasmid using FuGENE®HD Transfection Reagent (Promega) as per the manufacturer's protocol and the media was changed to DMEM supplemented with 10% FBS. The cells were then incubated at 28^{o}C for 24 hours. Fresh DMEM + 10% FBS was added to the cells before addition of steroid. The cells were incubated with steroid for 24 hours at 28^{o}C, and then harvested for assay. Measurement of transactivation was performed as above. The cells were incubated with 100μL passive lysis buffer for 15 minutes at room temperature. A 20μL aliquot was removed to a 96-well plate for assay, and relative light units were measured in a Perkin Elmer 2103 Envision Multilabile Reader. *Renilla* luciferase activity was determined to control for transfection efficiency.

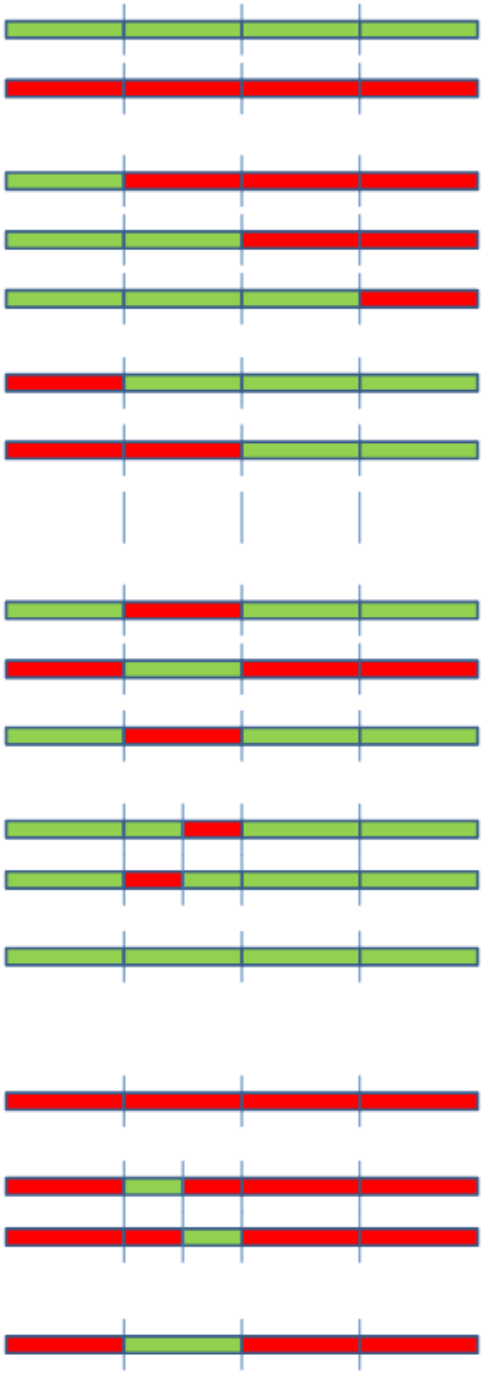
Western blot analysis. Western blotting with mouse monoclonal antibody MR1-18 (1:1000) (a gift from Prof. Celso Gomez-Sanchez) was used to demonstrate that the critical amino acid substitutions did not alter the stability of the receptor (*SI Appendix*, Fig. S4) The membrane was incubated with horseradish peroxidase-linked anti-mouse 2^{o} antibody (1:5000) (Dako, Carpinteria, CA, USA) for 60 minutes at room temperature. Antibody binding was visualized by chemiluminescence using the "Amersham ™ECL Plus Western Blotting Detection System" (GE Healthcare, UK). pCDNA3.1 vector alone and pRShGR (the equivalent glucocorticoid expression vector) were used as negative controls.

Statistical analyses. For the analysis of each wild-type, chimeric or mutanted MR the response relative to that of aldosterone for each construct was determined. Each data point represents the mean±SEM derived from two (discovery chimeras) or three independent transfection experiments, performed in triplicate on different days. Each figure represents an independent set of experiments providing independent validation where an MR construct examined in multiple figures/experiments. The mean values for each group were compared using a one-way ANOVA followed by Tukey's post hoc test with correction for multiple comparisons using Prism Version 7.0b (GraphPad Software Inc. San Diego, CA). The relative agonist/antagonist response for spironolactone or progesterone between MR is based on the percent agonist response, when compared to aldosterone, for that construct.

Molecular modelling. To achieve stabilisation of the MR LBD for crystallization, the approach used to obtain GR LBD crystals (**S1**) was applied to the equivalent amino acid in the hMR, such that the cysteine at position 808 in helix 5 was mutated to a serine (**15**,**31**). The original MR LBD crystals obtained by *Fagart et al.* (**17**) contained additional substitutions, Ser810Leu (**10**) and Cys910Ala; these mutations were used to stabilize the MR LBD to enable crystals to be obtained for x-ray crystallography. All x-ray crystal structures therefore maintain the MR LBD in a functionally active (agonist) form (**S2**). Thus, the antagonist conformation of the MR LBD is not represented by any of the currently published structures, irrespective of the ligand used (**S2**).
Models of native hMR and hMR-T870L LBD were generated using the MODELLER software (version 9.14) (**26**); the x-ray crystal structure of hMR with the highest resolution (PDB id: 4PF3) was used as the template for these models (**27**). Residues S808, L810 and V976 were converted to their native identities, C808, S810 and A976, respectively. The spironolactone ligand was docked into these models using VINA (**S3**) from within the YASARA program (version 17.4.17); the ligand-LBD complex structures with highest binding energy were used for subsequent MD simulations. The binding mode of spironolactone inside binding pocket was confirmed by comparison with the structure of hMR LBD in complex with spironolactone (PDB id: 2AB2) (**15**).

Molecular dynamics simulations. All MD simulations were performed using Gromacs software package (version 5.1.2) (28) with the GROMOS 54A7 United-Atom force field (29) and the G54A7FF United-Atom topology for spironolactone obtained using the ATB website (30). A steepest decent energy minimization was performed before short 100 ps NVT and NPT equilibration; production MD were performed from the previous NPT ensemble. The SPC water model was used to solvate systems, and all ionizable residues were assumed to be the standard state at pH 7; sodium and chloride ions were introduced by replacing water molecules to ensure the systems were electrically neutral and at a 0.1 M ionic strength. A cut-off of 10 Å in real space was applied to account for nonbonded interactions, and the particle-mesh Ewald method (S4) was applied for long-range electrostatics. The temperature was maintained at 300K using the velocity-rescale thermostat coupling method; protein and non-protein groups were treated independently. The time constant for coupling was set as 0.1 ps. The pressure of system was maintained at 1 bar with the Parrinello-Rahman barostat. The LINCS algorithm (S5) was used to constrain all bond lengths, allowing a 2 fs time step for MD integration. The Verlet grid cutoff-scheme was applied for neighbor searching with an update frequency of 10 steps. Periodic boundaries were applied in all directions.

Supplementary Information Figures



Designation	Transactivation spironolactone vs aldosterone
HHHH	5%
ZZZZ	75%
HZZZ	Not functional
HHZZ	17%
HHHZ	25%
ZHHH	6%
ZZHH	80%
HZHH	58%
ZHZZ	7%
HZHH	58%
HhzHH	22%
HzhHH	6%
HHHH	5%
ZZZZ	75%
ZhzZZ	44%
ZzhZZ	3%
ZHZZ	7%

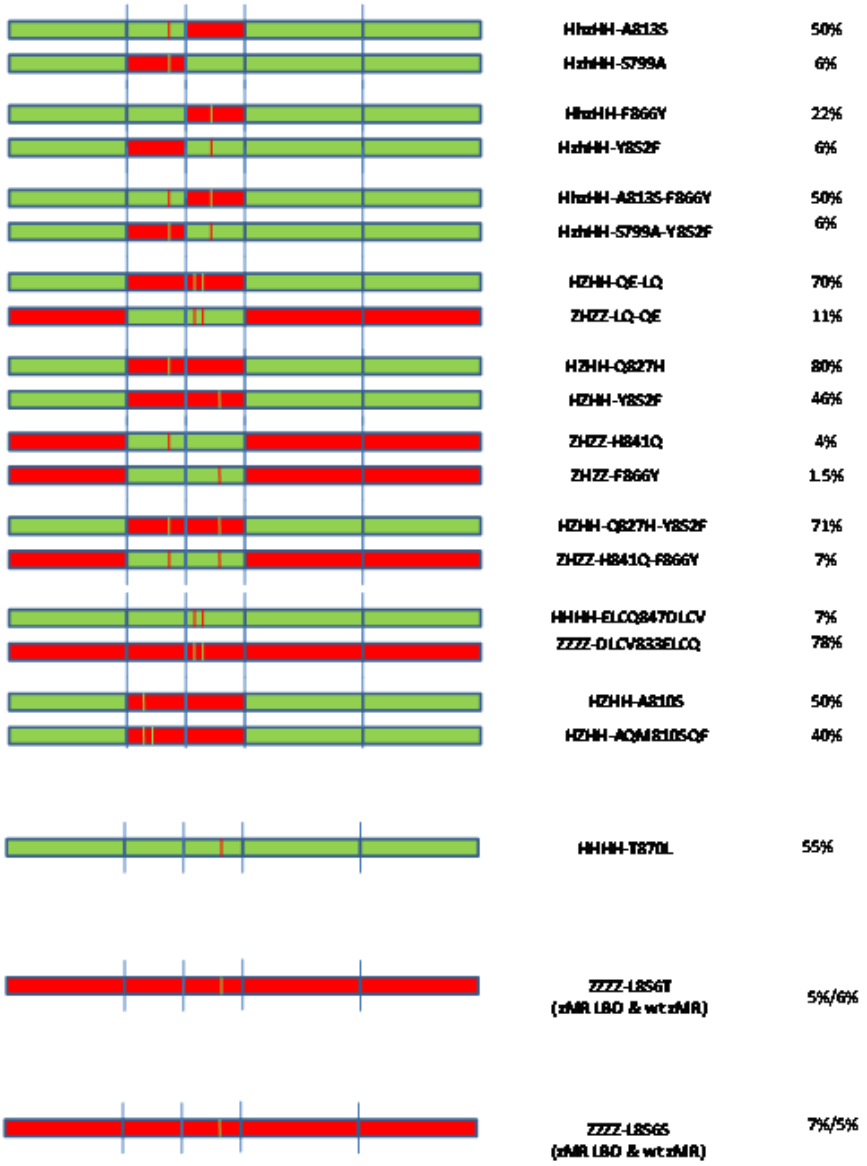


Fig. S1. hMR:zMR Chimeras and Mutations
Schematic representation of the series of hMR:zMR LBD chimeras created between amino acids 672-984 of the hMR LBD and the corresponding amino acids 658-970 of the zMR LBD. H is hMR sequence and Z is zMR sequence. Amino acids 785-880 of the hMR (zMR 771-866) correspond to the second of these four sections. Sub-chimeras were created within this second region in which hMR amino acids 785-843 were exchanged with zMR amino acids 771-829 in the zMR LBD and amino acids 844-880 from the hMR were exchanged with amino acids 830-866 in the zMR. The primers used to create the site-drected mutagenesis are in Table S1.

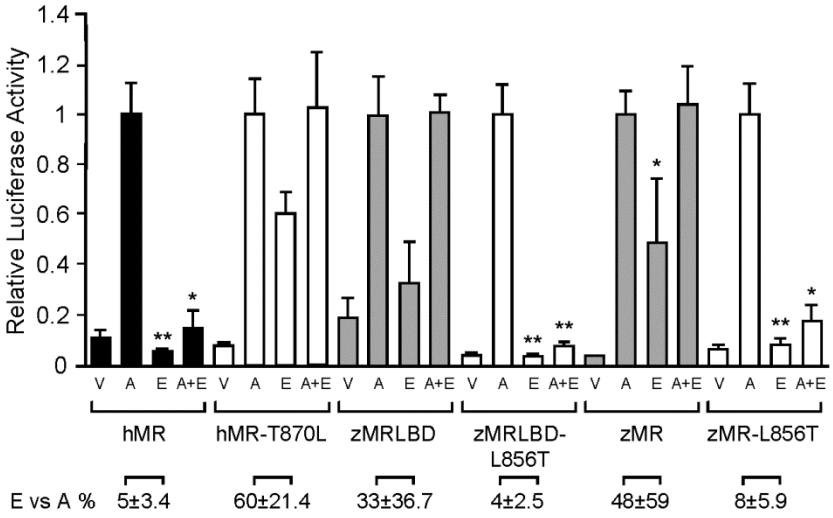


Fig. S2. Transactivation responses of hMR 870 leucine and zMR 856 threonine to aldosterone and eplerenone.

The response of the wildtype hMR (hMR) and hMR threonine 870 leucine (hMR-T870L); and both zMRLBD and wildtype zMR (zMR) with leucine 856 threonine (zMR-L856T and zMRLBD-L856T, respectively) to 1nM aldosterone (A) and 10μM eplerenone (E) was examined as described for Figures 2 and 4. Each data point represents the mean ± SEM derived from two independent experiments with all treatment groups (A, E and A+E) being significantly greater ($p < 0.05$) than vehicle alone except for hMR, zMRLBD, zMRLBD-L856T and zMR-L856T in response to spironolactone; and hMR and zMR-L856T in response to both (A+E). Eplerenone and eplerenone plus aldosterone are less than aldosterone alone (*$p<0.01$; **$p<0.001$) where indicated. The activation with eplerenone alone as a percentage of that for aldosterone (E vs A) for each MR below the graph; the relative response to spironolactone significantly differs between the hMR and hMR-T870L ($p < 0.001$) but does not achieve significance for zMR LBD ($p=0.07$) and zMR ($p=0.097$) when compared to their corresponding mutants.

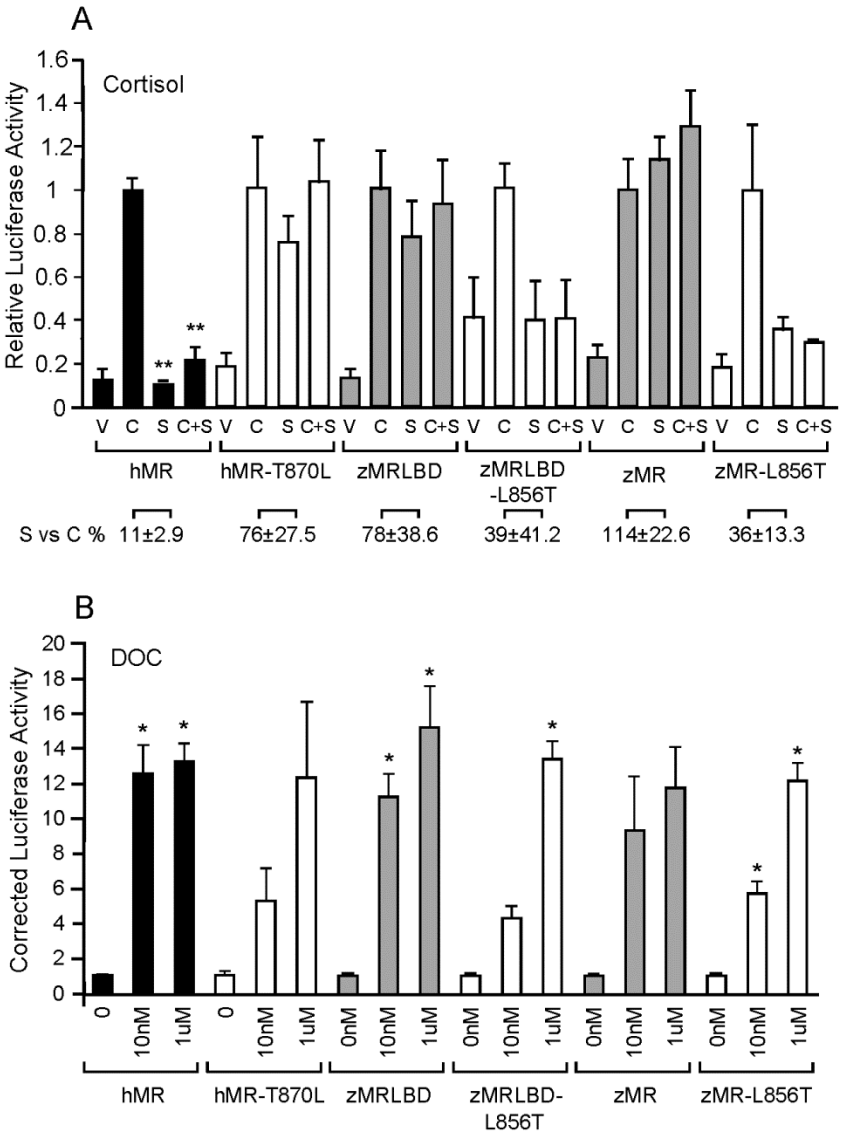


Fig. S3: Transactivation response of zMR 856 threonine to cortisol and spironolactone and DOC.

A) Response of wildtype hMR (hMR) and hMR threonine 870 leucine (hMR-T870L); and both wildtype zMR (zMR) and zMRLBD alone and containing leucine 856 threonine (zMR-L856T and zMRLBD-L856T, respectively) to 10nM cortisol (C), 1 μM spironolactone (S) and both (C+S) was examined as described for Figure 2. Each data point represents the mean ± SEM derived from two independent experiments with all treatment groups (C, S and C+S) being significantly greater ($p < 0.05$) than vehicle alone except for hMR and zMRLBD-L856T for spironolactone alone and zMR-L856T with both. Spironolactone and spironolactone plus cortisol are less than cortisol alone (** $p<0.001$) where indicated. The activation with spironolactone alone as a percentage of that for cortisol (S vs C) for each MR is shown below the graph; the relative response to spironolactone significantly differs ($p<0.01$) between the intact wildtype MR and the corresponding mutant MR LBD. B) Response of wildtype hMR (hMR) and hMR threonine 870 leucine (hMR-T870L); and both wildtype zMR (zMR) and zMRLBD alone and containing leucine 856 threonine (zMR-L856T and zMRLBD-L856T, respectively) to vehicle (0), 10nM DOC and 1μM DOC. Each data point

represents the mean ± SEM derived from two independent experiments with all treatment groups being significantly greater than vehicle alone indicated (* p< 0.05).

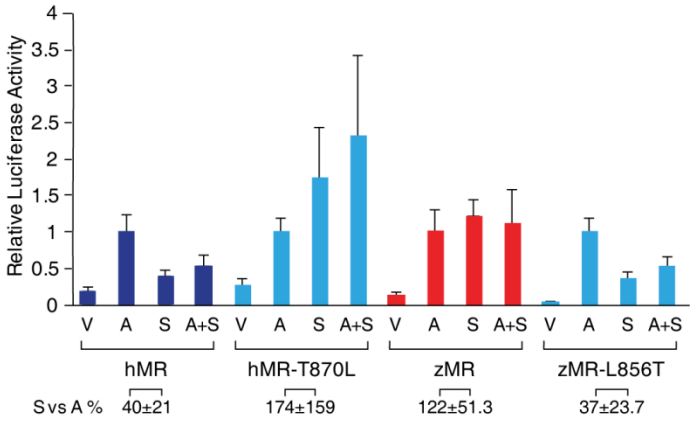


Fig. S4: Transactivation responses of hMR, hMR 870 leucine, zMR and zMR 856 threonine to spironolactone in ZF4 cells.
The approach is described for Fig. 1*E* with each data point representing the mean ± SEM derived from three independent experiments. The relative response to S vs A, significantly differs between the intact MR LBD and the corresponding mutant MR LBD (p< 0.05).

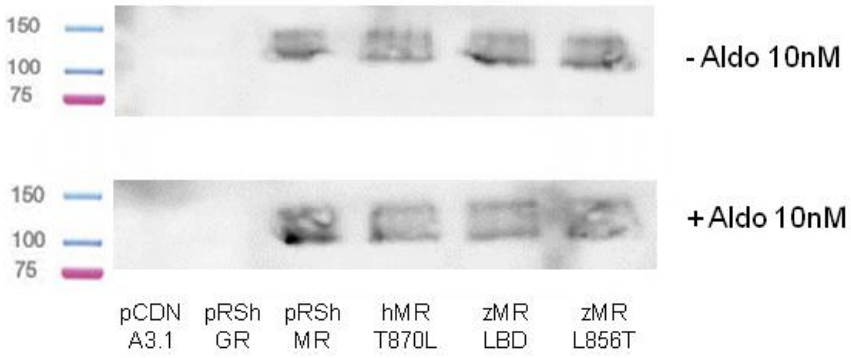


Fig. S5. Western blot analysis
Western blot analysis analyses of the hMR:zMR chimera using a N-terminal directed MR antibody. The autoradiogram shows the result for the MR with or without aldosterone in the presence of empty vector (pCDNA3.1), the glucocorticoid receptor (pRShGR), pRShMR, pRShMR with the threonine 870 leucine substitution (hMR T870L), the zebra fish MR LBD alone (zMRLBD) and with the leucine 856 threonine substitution (zMR L856T).
The molecular markers are at the left.

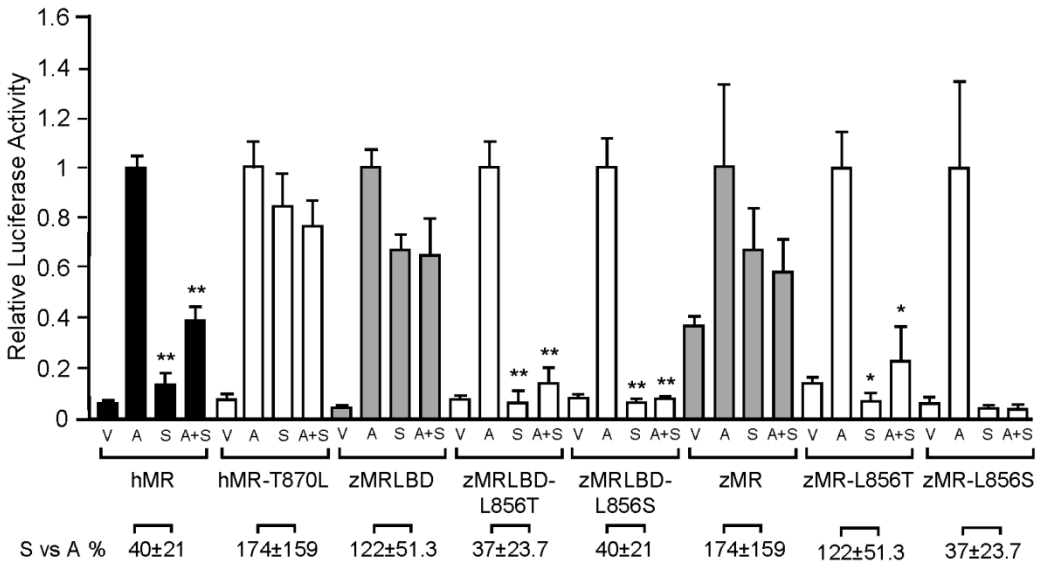


Fig. S6. Transactivation responses of zMR 856 serine (zMR-L856S and zMRLBD-L856S) to aldosterone and spironolactone.

The response of wildtype hMR (hMR) and hMR threonine 870 leucine (hMR-T870L); and both zMRLBD and wildtype zMR (zMR) with leucine 856 threonine (zMR-L856T and zMRLBD-L856T respectively); and both zMRLBD and wildtype zMR (zMR) with leucine 856 serine (zMR-L856S and zMRLBD-L856S respectively) to 10nM aldosterone (A) and 1μM spironolactone (S) was examined as described for Figure 2. Each data point represents the mean ± SEM derived from two independent experiments with all treatment groups (A, S and A+S) being significantly greater ($p < 0.05$) than vehicle alone ($p < 0.05$) except for hMR and zMR wildtype and the zMR containing either threonine or leucine at position in response to spironolactone; and zMR wildtype and the zMR containing either threonine or leucine at position in response to both (A+S). Spironolactone and spironolactone plus aldosterone are less than aldosterone alone (*$p<0.05$) where indicated. The activation with spironolactone alone as a percentage of that for aldosterone (S vs A) for each MR is shown below the graph; the relative response to spironolactone significantly differs between the intact MR LBD and the corresponding mutant MR LBD ($p < 0.05$).

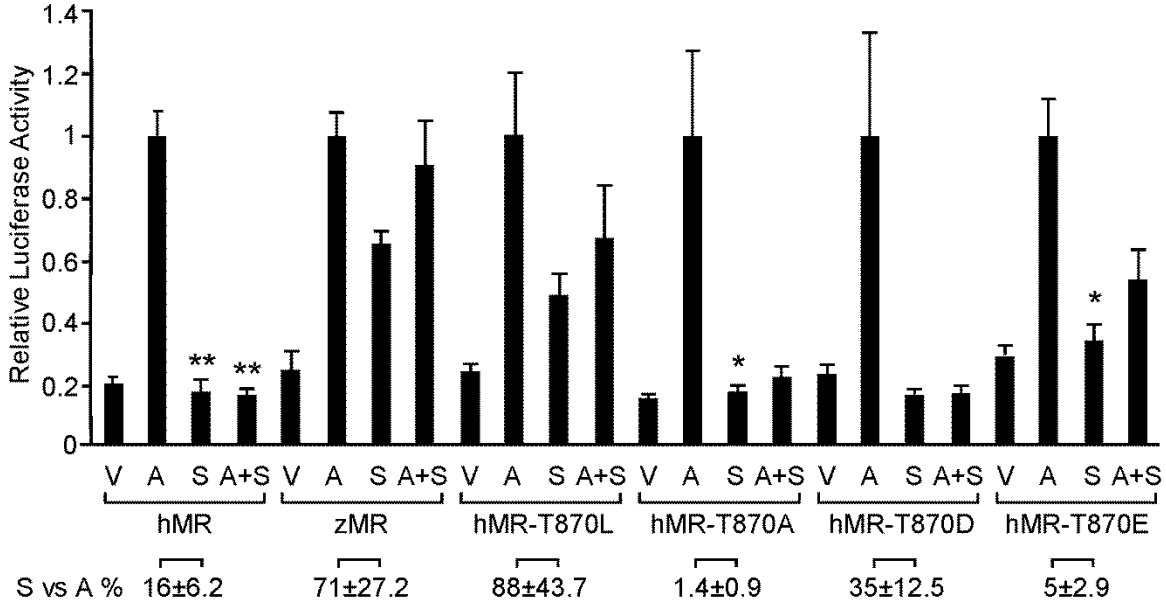


Fig. S7: Transactivation responses of zMR 856 aspartate, glutamate or alanine to aldosterone and spironolactone.

The response of wildtype hMR (hMR) and wildtype zMR (zMR) together with hMR threonine 870 leucine (hMR-T870L), hMR threonine 870 alanine (hMR-T870A) hMR threonine 870 aspartate (hMR-T870D)and hMR threonine 870 glutamate (hMR-T870E) to 10nM aldosterone (A) and 1μM spironolactone (S) was examined as described for Figure 2. Each data point represents the mean ± SEM derived from two independent experiments. Aldosterone treatment is significantly different ($p<0.05$) from vehicle for MR with all treatment groups. Spironolactone or spironolactone plus aldosterone is significantly greater ($p< 0.05$) from vehicle alone for hMR containing leucine or for zMR. Spironolactone and spironolactone plus aldosterone are less than aldosterone alone (*$p<0.05$) where indicated. The activation with spironolactone alone as a percentage of that for aldosterone (S vs A) for each MR is shown below the graph.

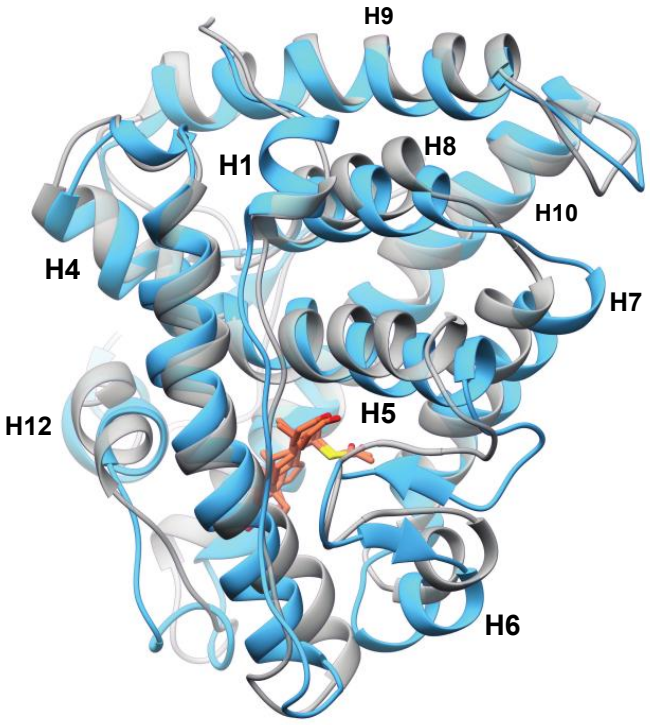


Fig. S8. Superposition of native hMR and hMR-T870L structures after 2 μs.
Native hMR LBD is coloured grey and hMR-T870L is coloured cyan. Ligand carbon atoms are coloured in coral, oxygen in red and sulfur in yellow. Helices are labelled H1 through H12.

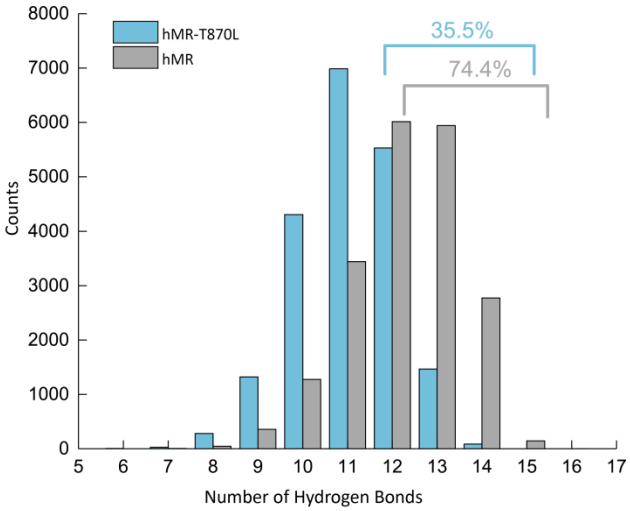


Fig. S9. Distribution of intra-helix hydrogen bonds throughout the 2 μs trajectory for native and T870L hMR LBD.
For native hMR, 74.4% of the structures sampled throughout the MD simulation included 12-or-more hydrogen bonds in helix-8, whereas for hMR-T870L only 35.5% of structures sampled included the same number of hydrogen bonds. The hydrogen bond between the hydroxyl group of T870 and backbone carbonyl of F866 contribute to stabilizing helix-8 in native hMR. Native hMR LBD is coloured grey and hMR-T870L is coloured cyan.

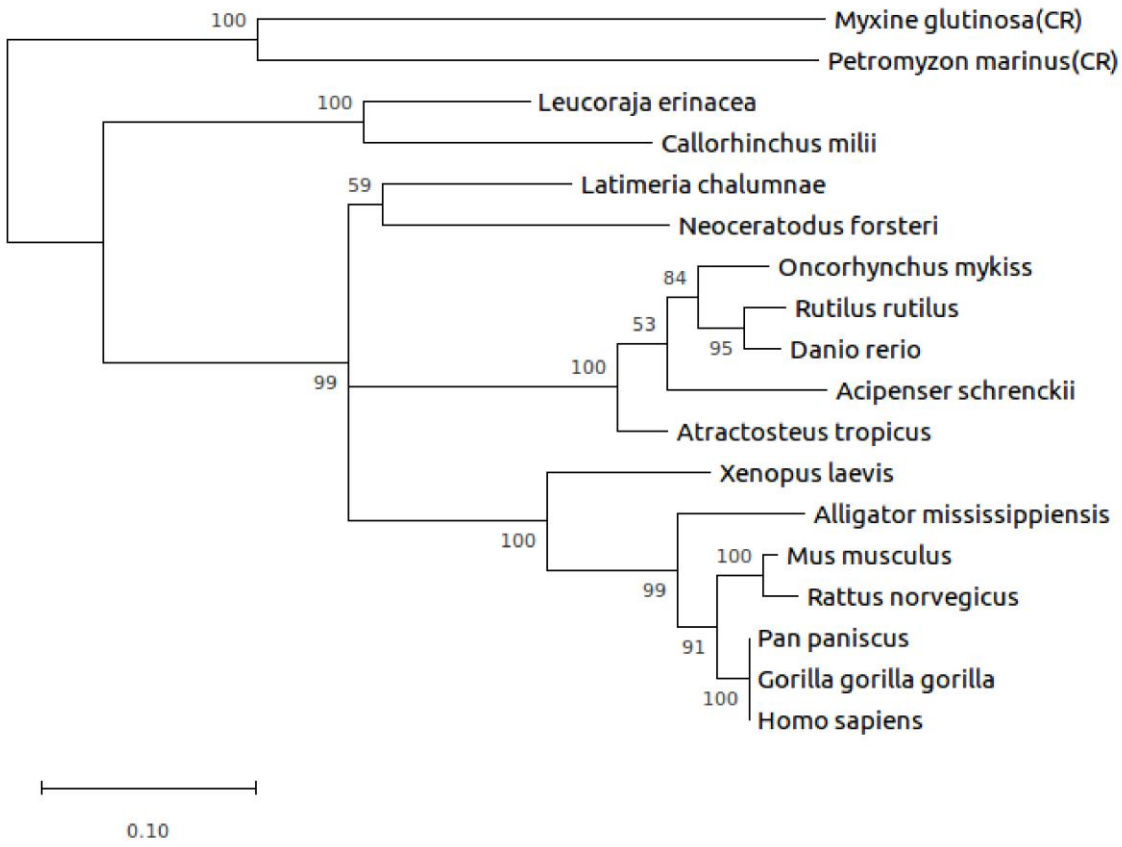


Fig. S10. Molecular Phylogenetic analysis of the MR by Maximum Likelihood method.
The evolutionary history of vertebrate animals from sea to land was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (**S6**). The tree with the highest log likelihood (-2872.00) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.4245)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 19.60% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 16 MR LBD sequences and 2 CR LBD sequences. The LBD sequences were collected from the NCBI Protein database, multiple sequences were aligned via ClustalW algorithm (**S7**). There were total of 250 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (**S8**).

References

S1. Bledsoe RK, Montana VG, Stanley TB, Delves CJ, Apolito CJ, McKee DD, Consler TG, Parks DJ, Stewart EL, Willson TM, Lambert MH, Moore JT, Pearce KH, Xu HE (2002) Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. *Cell* **110**:93-105.

S2. Huyet J, Pinon GM, Fay MR, Rafestin-Oblin M-E, Fagart J (2012) Structural determinants of ligand binding to the mineralocorticoid receptor. *Mol Cell Endocrinol* **350**:187-95.

S3. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**:455–461.

S4. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* **103**:8577-8593.

S5. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* **18:**1463-1472.

S6. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **8**:275-282.

S7. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD,Gibson TJ, Higgins DG, (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947-2948.

S8. Kumar S, Stecher G, Li M, Knyaz C, Tamura K, (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evo.* **35**:1547-1549.