Appendix C

Altering the specificity of an androgen receptor

This project was carried out in collaboration with Prof. Robert Fletterick's lab at the University of California, San Francisco. The computational components of the project described here were carried out in the Mayo lab. All experimental work was carried out by Leslie Cruz in the Fletterick lab.

Abstract

Androgen receptors (AR) are nuclear hormone receptors that play a major role in sexual development. To study the role of AR on sexually stereotyped behavior, a mutant AR/ligand pair is needed that is orthogonal to the wild-type AR/androgen system. This variant AR could then be selectively activated in the presence of the wild-type AR and wild-type ligands. Here, we attempted to use our computational protein design methods to design an AR that can be activated by the nonnatural ligand 19PT. Our results so far indicate that none of our designed ARs are activated by 19PT.

Introduction

Some behavioral differences between the two sexes of a species are controlled on the molecular level. Male mice, for example, demonstrate aggressiveness towards other mice and are territorial, whereas female mice do not exhibit these behaviors.[1] These sexually dimorphic behaviors have prompted investigations into biochemical differences in the brains of male and female mice. However, few examples of molecular differences have been identified so far. One major difference has been observed in the distribution of androgen receptors (AR) in the hypothalamus of mice; male mice tend to have a larger number of highly localized neurons expressing AR, whereas females have fewer, more disperse neurons expressing AR.[1]

AR is a 110 kD nuclear hormone receptor, that is sequestered by a heat shock protein (HSP) in the cytoplasm. Upon binding dihydrotestosterone (DHT), a metabolic derivative of testosterone (TES), AR undergoes a conformational change that allows its translocation into the nucleus. Once in the nucleus, it homodimerizes and acts as a DNA transcription activator for genes that control sexual development and the maintenance of skeletal and muscular systems (Figure C-1).[2,3] While these roles of AR are widely accepted, less is known about the role of AR in the brain and its effect on sexually stereotypical behavior.[1]

Our goal here was to create a system that can facilitate the study of the role of AR in sex-differentiated behavior in a mouse model. To this end, we attempted to design an AR that can be activated by the nonnatural TES-analog 19PT (Figure C-2). The design of a novel AR-ligand pair is the first step to the design of a "magic pair" that is orthogonal to the wild-type AR-DHT system. The expression of this mutant AR in the

brains of mice would allow manipulation of mutant AR activity in a site-specific manner without affecting the endocrine system in the rest of the animal.

The strategy for redesigning the AR active site to accommodate 19PT is similar to that for designing enzyme active sites described in Chapters I-IV. As before, we used the computational protein design software ORBIT.[4] However, instead of designing around a reaction transition state (TS), we optimize the active site residues for binding to the ground-state ligand structure. Important binding contacts, such as hydrogen bonds, can be enforced using geometric constraints as in the enzyme design calculations. In addition, some rotational and translational freedom is given to the ligand during the design calculation and the internal flexibility of the ligand is modeled with canonical torsions of rotatable bonds, as in Appendix A.

Methods

Scaffold selection

Because it had been used for past analyses by the Fletterick lab, the 1.7 Å structure of AR from *Pan troglodytes* bound to DHT was chosen as the scaffold for the design.[5]

19PT structure and rotamer design

The DHT structure found in the crystal structure from Hur et al.[5] (PDB: 1T7T) was modified to produce the structure of 19PT using the molecular modeling program BIOGRAF.[6] Starting with the DHT structure, hydrogens and the 19-butyl group were added and the entire structure was minimized for 50 steps. Charges were calculated

using an electrostatic potential fit, methanol solvation, and the hybrid density functional B3LYP as implemented by the Jaguar 5.5 software package and a 6-31G** basis set.[7] Rotamers were created using canonical torsions (60°, 180°, and 300°) for each of the four rotatable bonds, resulting in 81 rotamers, which were then minimized for 250 steps using BIOGRAF (Figure C-3).[6]

Androgen receptor design for 19PT binding

The initial position of the 19PT ligand was determined by overlaying with the DHT structure in the 1T7T crystal structure.[5] The ligand was allowed to rotate ± 5° in x, y, and z in 5° steps and translate ± 0.2 Å in x, y, and z in 0.2 Å steps. Three positions (W741, M742 and M745) were chosen as design positions because of their proximity to the expected position of the 19PT butyl group based on its initial position (Figure C-4). All residues within 6 Å of DHT in the 1T7T structure (701, 704, 705, 707, 746, 749, 752, 746 768, 780, 784, 787, 873, 876, 877, 880, 889, 891, 895, 899) were allowed to change conformation but not identity.

As in Chapter III, Lazaridis-Karplus occlusion-based solvation was applied with scale factors of 1.0 for nonpolar burial and nonpolar exposure and a scale factor of 0.6 for polar burial (see Materials and Methods).[8] Other standard ORBIT parameters were applied as in Lassila *et al.* and a backbone-independent conformer library was used to represent sidechain flexibility[9]. Loose geometric constraints were applied to preserve hydrogen bonds that are present in the wild-type crystal structure. An Asp and Thr residue was required to be between 2.5 and 3.2 Å of one of the 19PT hydroxyls and an arginine contact was required between 4.0 and 5.0 Å of the other hydroxyl (Figure C-5).

Transcriptional activation assay

The activation of AR variants was monitored using a luciferase reporter assay in HeLa cells as described by Bohl *et al.*[10] In these experiments, the AR ligand binding domain is fused to the transcriptional activator Gal4 (Gal4-AR). Ligands were added to the culture after a day of expression and chemiluminescence was used to determine AR activation. Assays were carried out with 10 μM 19PT and 100 nM DHT. Percent activity was calculated based on DHT activity in Gal4-AR set at 100%.

Results and Discussion

The designs from the ORBIT calculations are summarized in Table C-1. In general, the sequences are hydrophobic and relocate the bulk in the active site from position 745 to position 715 to make room for the 19PT butyl group. The sequences resulting from the calculations fall into five groups, which can be separated based on the residue chosen at position 741. Groups 1-4 each contain 3 sequences that all have an alanine at position 745. All three sequences in each of these groups are structurally identical except for the residue at 715, which is either leucine, methionine, or isoleucine. These four groups are distinguished by the residue at position 741.

Group 1 has a tyrosine at position 741, whose hydrogen bond contact is unfulfilled. The 19PT structure in these designs does not have an extended butyl group. Group 2 designs are identical in structure to group 1, except that a phenylalanine is substituted for tyrosine at position 741 (Figure C-6A). Again, the butyl group of 19PT is

in a kinked conformation that may not be energetically favorable. Group 3 designs have a W715H mutation. In these designs, the histidine overlays well with the indole of the tryptophan in the wild-type crystal structure and the butyl of 19PT is in a fully extended conformation (Figure C-6B). Group 4 has a leucine at position 741, which causes the butyl group to adopt the same kinked conformation seen in group 2 (Figure C-6C). In contrast to the other sequences, group 5 has only a single sequence with a single W741A mutation and the 19PT butyl group is in a fully extended conformation (Figure C-6D).

These five groups of designs were submitted to the Fletterick lab for testing and a large number of variants were made. Only one of the variants (W741A) shows wild type-like DHT activation (Gal-AR) in the transcription activation assay and none of the AR variants that we designed so far shows significant activation by 19PT (Figure C-8). The loss of DHT activation in most of the designs indicates some kind of perturbation in the active site or a prevention of the conformational change necessary for activation. One of the limitations of computational protein design is that we can only design for ligand binding and not for AR activation. Unfortunately, in the case of AR, ligand binding is necessary but not sufficient for activation. To evaluate the success of the computational protein design process in this case, we would need to assay for binding directly. Currently, no binding assay exists for AR beyond crystallography, which is extremely low-throughput and not a reliable assay for binding due to the high potential for false negatives.

In addition, conformational changes of the backbone upon mutation of the binding pocket residues cannot be ruled out. In the crystal structure of AR/M745A, the backbone of the binding pocket is observed to shift slightly, changing the conformation of some of

the residues significantly, including W741 and N705 (Figure C-7).[11] Future design calculations could account for this backbone flexibility by using molecular dynamics or by including multiple backbone conformations during the design calculation.[12,13]

Conclusions

So far, we have not been able to use our computational protein design methods to design a mutant AR that can be selectively activated by the nonnatural ligand 19PT. However, without specific structural or ligand-binding data, it remains unclear if our methods failed or if this system is simply not amenable to computational design due to subtleties in the receptor structure-function relationship. This project is ongoing with experimental work in the Fletterick lab. Future computational studies of this system with molecular dynamics simulations may help us understand the effect of mutations on the activity of the receptor, and the incorporation of this information into future designs using multi-state design may aid us in creating active AR variants.

References

1. Shah, N. M.; Pisapia, D. J.; Maniatis, S.; Mendelsohn, M. M.; Nemes, A.; Axel, R., Visualizing sexual dimorphism in the brain. *Neuron* **2004**, *43*, 313-319.

2. Mooradian, A. D.; Morley, J. E.; Korenman, S. G., Biological actions of androgens. *Endocr. Rev.* **1987**, *8*, 1-28.

3. Keller, E. T.; Ershler, W. B.; Chang, C., The androgen receptor: a mediator of diverse responses. *Frontiers in Bioscience* **1996**, *1*, d59-71.

4. Street, A. G.; Mayo, S. L., Computational protein design. *Structure* **1999**, *7*, 105-109.

5. Hur, E.; Pfaff, S. J.; Payne, E. S.; Gron, H.; Buehrer, B. M.; Fletterick, R. J., Recognition and accommodation at the androgen receptor coactivator binding interface. *PLoS Biol.* **2004**, *2*, 1303-1312.

6. BIOGRAF version 3.21, Molecular Simulations, Inc., Burlington, MA, **1992**.

7. Jaguar 5.5, Schrodinger, L.L.C., Portland, OR, **1991-2003**.

8. Lazaridis, T.; Karplus, M., Effective energy function for proteins in solution. *Proteins* **1999**, *35*, 133-152.

9. Lassila, J. K.; Privett, H. K.; Allen, B. D.; Mayo, S. L., Combinatorial methods for small-molecule placement in computational enzyme design. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16710-16715.

10. Bohl, C.; Miller, D. D.; Chen, J.; Bell, C. E.; Dalton, J. T., Structural basis for accommodation of nonsteroidal ligands in the androgen receptor. *J. Biol. Chem.* **2005**, *280*, 37747-37754.

11. Cruz, L.; Fletterick, R. J., (*Unpublished results*). **2007**.

12. Allen, B. D. *Development and validation of optimization methods for the design of protein sequences and combinatorial libraries*. California Institute of Technology: Pasadena, CA, **2009**.

13. Nisthal, A.; Allen, B. D., (*Manuscript in preparation*). **2009**.

Table C-1. AR-19PT design summary. The designs were separated into five groups based on their mutations at positions 741 and 745. Group 1 (blue) has W741Y and M745A mutations. Group 2 (orange) has W741F and M745A mutations. Group 3 (white) has W741H and M745A mutations. Group 4 (purple) has W741L and M745A mutations. Group 5 (yellow) has only a W741 mutation.

name	715	741	745	ORBIT energy (kcal/mol)	comments
wt	V	W	M		
1	L	Y	A	-216.9	Y742 makes no H-bond, butyl not fully extended
2	M	Y	A	-216.8	
3	I	Y	A	-216.8	
4	L	F	A	-216.0	butyl not fully extended, but takes up most of empty space
5	M	F	A	-215.9	
6	I	F	A	-215.9	
7	L	H	A	-215.5	His overlays well with indole of W741
8	M	H	A	-215.4	
9	I	H	A	-215.4	
11	L	L	A	-215.1	W241L has been made previously and does not preturb structure
12	M	L	A	-215.1	
13	I	L	A	-215.1	
14	V	A	M	-212.6	single mutation

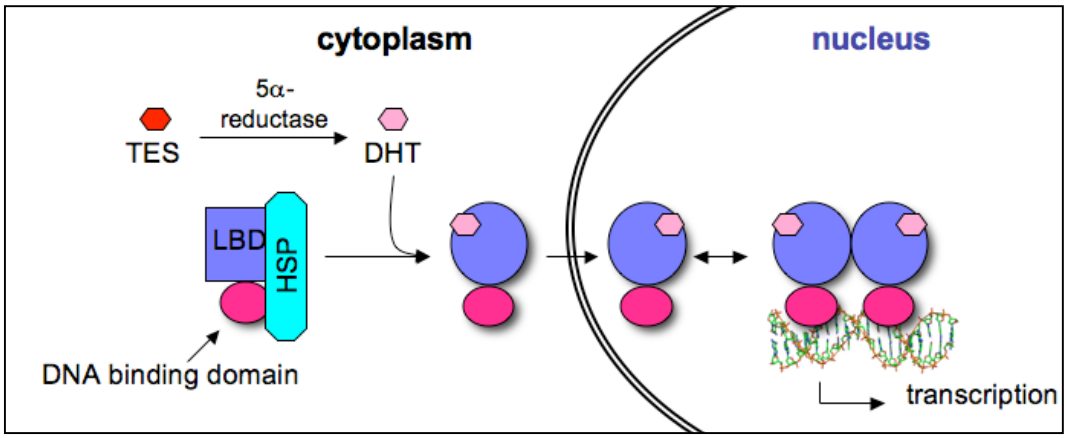
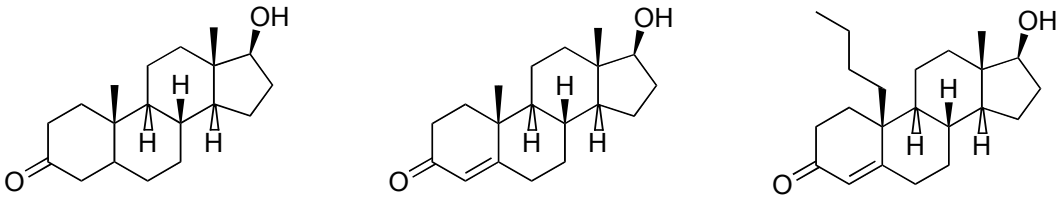


Figure C-1. Activation mechanism of AR.



dihydrotestosterone **testesterone** **19PT**

Figure C-2. Chemical structures of androgens of interest.

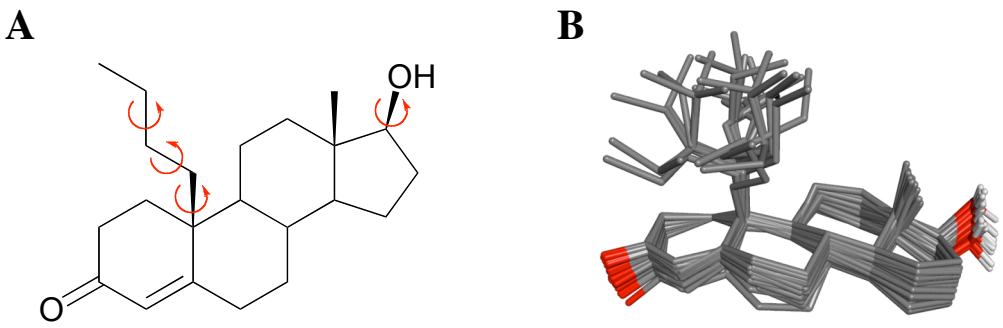


Figure C-3. Rotamers of 19PT. **(A)** Chemical structure of 19PT with red arrows indicating the positions of rotatable bonds. **(B)** Overlay of the 81 minimized 19PT rotamers used in the calculations.

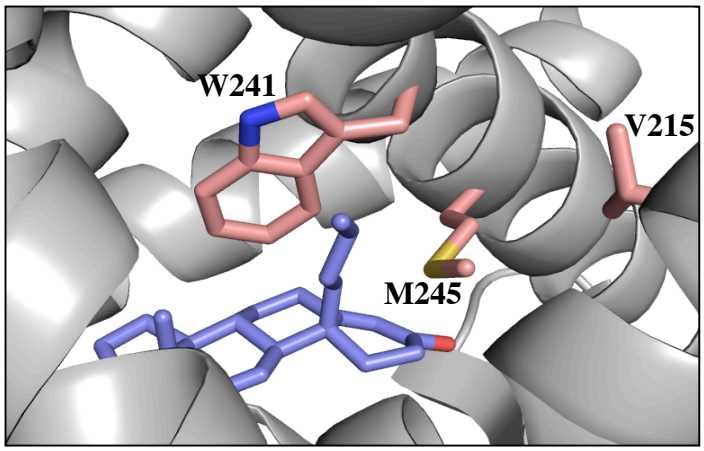


Figure C-4. Design positions in the active site of AR. Sidechains that were designed are indicated and a model of the 19PT ligand is shown in purple.

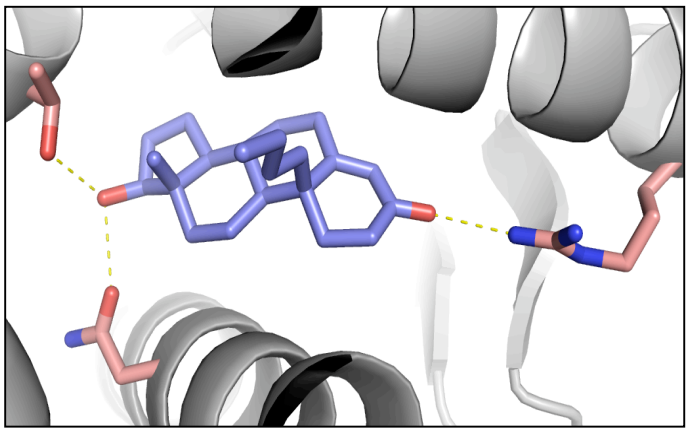


Figure C-5. Wild-type hydrogen bonds to 19PT. These hydrogen bond contacts were enforced through a simple geometry pruning step. The 19PT model is shown in purple.

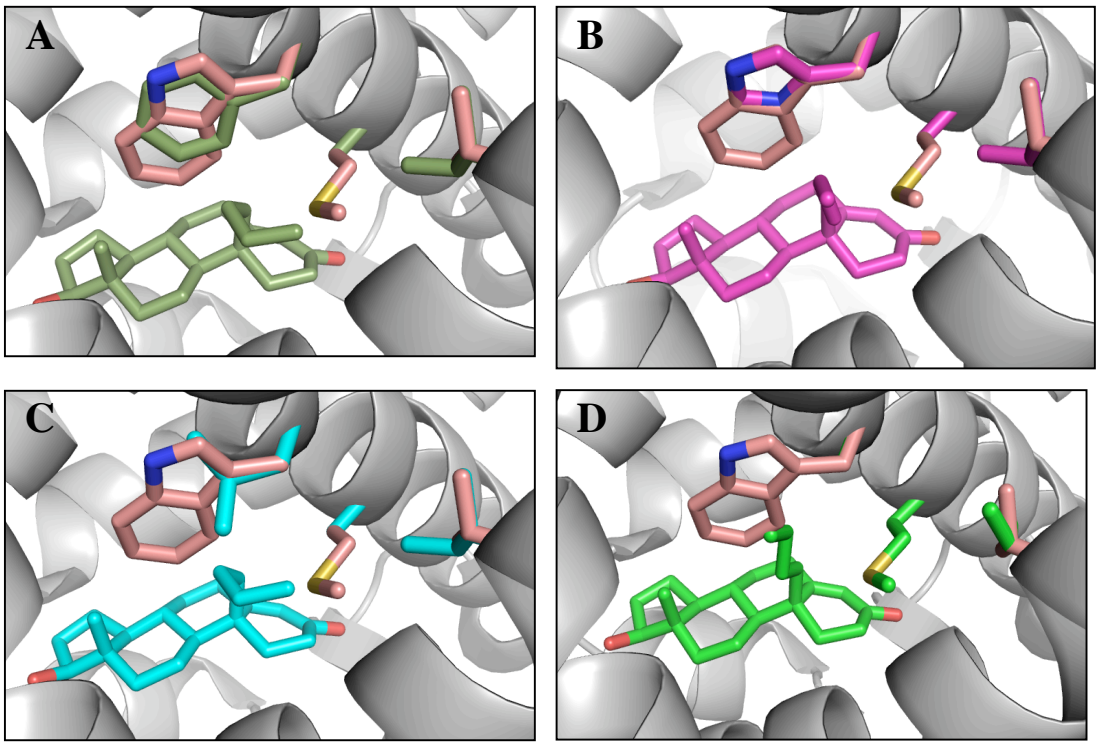


Figure C-6. ORBIT designs for AR binding of 19PT. Each design is shown overlaid with the wild-type AR crystal structure, shown in pink.[5] **(A)** AR/V215I/W241F/M245A design is shown in green. **(B)** AR/V215I/W241H/M245A is shown in magenta. **(C)** AR/V215I/W241L/M245A is shown in cyan. **(D)** W241A is shown in green.

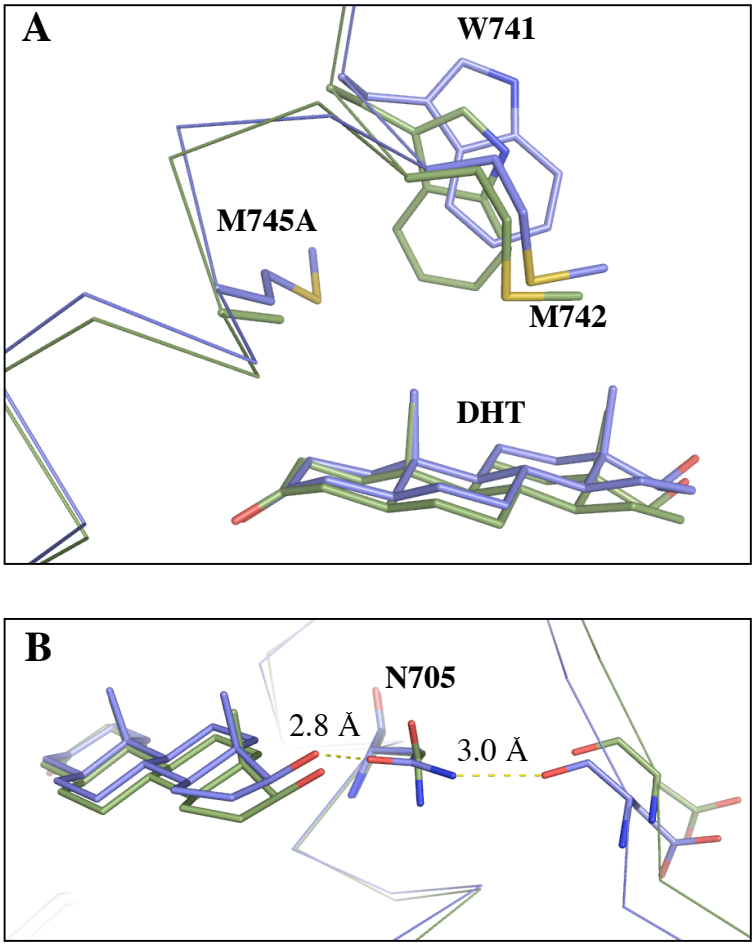


Figure C-7. Structural differences in the AR crystal structure caused by M745A. The structure of wild-type AR[5] is shown in blue and the structure of AR/M745A[11] is shown in green. DHT was bound in both structures. **(A)** Structural differences of W741 and M742. **(B)** Structural differences of N705.

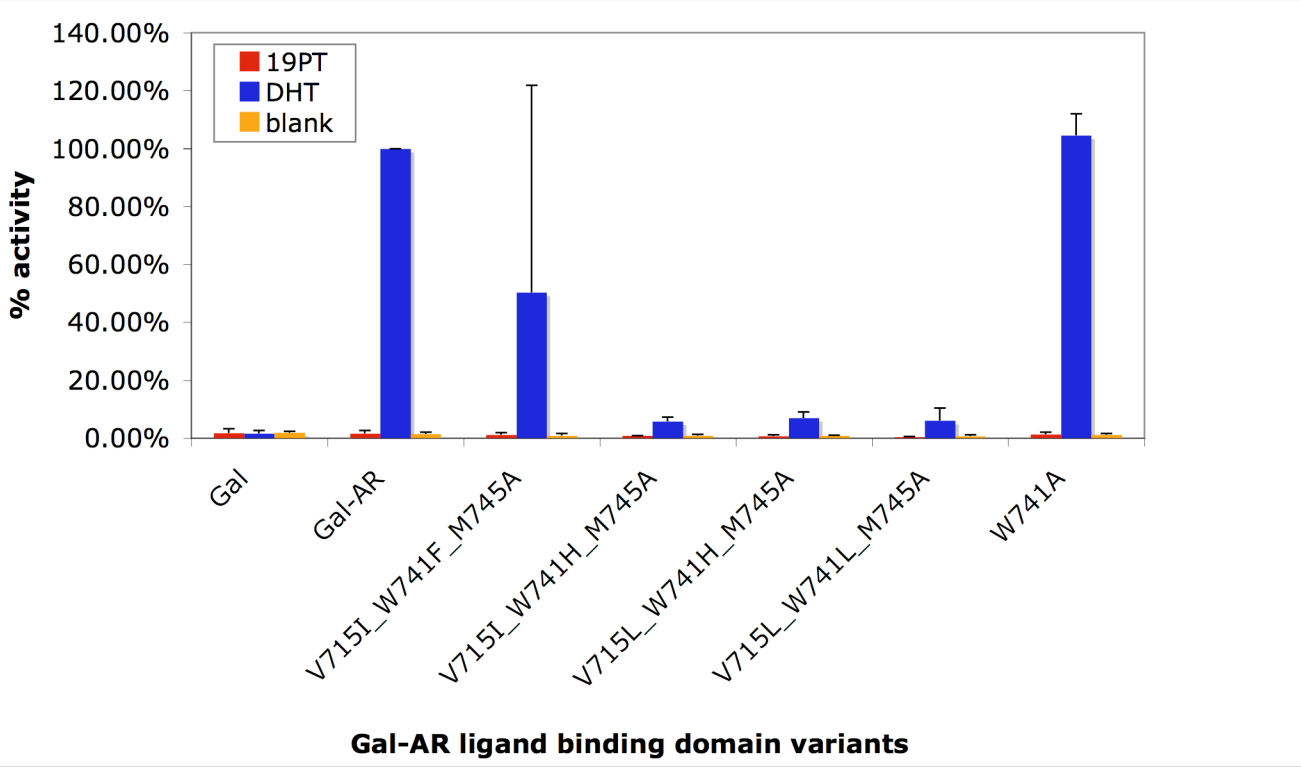


Figure C-8. Transcriptional activation assay. Percent activity is calculated based on the DHT activity in Gal-AR being 100%.