JCTC Journal of Chemical Theory and Computation
Letter
pubs.acs.org/JCTC

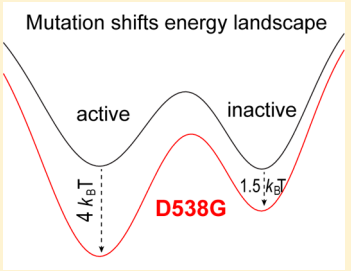
# A Newfound Cancer-Activating Mutation Reshapes the Energy Landscape of Estrogen-Binding Domain

Wei Huang, Krishnakumar M. Ravikumar, and Sichun Yang*

Center for Proteomics and Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106-4988, United States

Supporting Information

**ABSTRACT:** The ligand-binding domain (LBD) of an estrogen receptor undergoes a large conformational switching from an inactive to active state in response to hormone stimuli. Very recently, a novel D538G mutant has been identified to be active in advanced breast cancer tumors. Here, we ask if molecular simulations can provide insight on its mechanistic impact on the receptor's activation status. It has been challenging for *ab initio* modeling to identify two distinct conformations of a single amino acid sequence as large as that of the LBD. Using a coarse-grained (CG) model, we are able to correctly reproduce this LBD conformational switching. Furthermore, we found that the D538G mutation reshapes the energy landscape by stabilizing both active and inactive conformations, but preferring the active by 1.5 kcal/mol. This observation is consistent with the concept of a mutation-shifting landscape and provides a structural explanation for the oncogenic D538G mutation at the detailed conformational level.




Conformation changes are critical for many proteins to function in response to external stimuli. How to model such a large transition is challenging for *ab initio* modeling. Here, we show that the use of an effective coarse-grained energy model is able to correctly predict the structural transition between two well-defined conformations of the ligand-binding domain (LBD) of estrogen receptor alpha (ER$\alpha$). The ability of capturing this large-scale transition not only validates this coarse-grained energy model itself but also provides a conformation-driven explanation of a recently observed clinical ER mutation that reshapes the entire energy landscape.

The ER ligand-binding domain is a remarkable example of conformational switching, where a C-terminal $\alpha$-helix H12 is repositioned from an inactive to active conformation.[1−3] The inactive conformation can be stabilized by therapeutic inhibitors such as raloxifene, so H12 is in a helix-bundle forming conformation (marked in blue in Figure 1A) that can block the receptor from recruiting coactivators for transcriptional activation. Once activated by estrogen binding, H12 is switched to a completely different orientation covering the estrogen-binding pocket in the active conformation (marked in red in Figure 1A). These two well-defined conformations—active and inactive—have been revealed extensively by crystallographic studies.[1,2,4] Furthermore, several groups have recently reported a novel mutation of D538G (at the beginning of H12), as found in advanced breast cancer patients.[5−7] Structure modeling of this mutation has suggested a conformation-driven mechanism, where the mutation stabilizes the active conformation.[5,6] While this information is highly informative, a key question remains, i.e., whether the mutation affects the active state only or sculpts the entire landscape, including both inactive and active states. To answer this question from a broad landscape perspective, the modeling of such a large-scale H12 transition can be quite challenging for atomic-level simulations. It is even more so given the large conformation space involved to observe frequent switching events.

Here, an effective coarse-grained model is used to simulate the LBD conformational dynamics at the residue level. One simplification comes from the partitioning of the LBD into three parts: ΔLBD ($S_{301}$−$K_{529}$), linker ($C_{530}$−$Y_{537}$), and H12 ($D_{538}$−$R_{548}$) (Figure S1A), in accordance with local flexibility observed in hydrogen−deuterium exchange measurements.[8] A structure-based Gō-type energy function is employed to describe internal interactions within ΔLBD and H12 separately, while the interactions between ΔLBD and H12 are modeled using a predictive coarse-grained model (as opposed to a Gō-type function; see Supporting Information). That way, no prior structural information was taken from solved crystal structures to model the ΔLBD-H12 interactions. In fact, only the protein sequence was used to choose from and set the physical parameters for the electrostatic and hydrophobic interactions, which are mostly pairwise terms between all 20 amino acids (Table S1). Note that this energy function was initially developed for simulating protein−protein interactions[9] but has directly been transferred here to model the H12 switching. For the connecting linker, a chain of harmonic springs was used to model its intrinsic flexibility. It should also be noted that while the bound ligands were not explicitly modeled, the ligand-binding pocket was implicitly mimicked by imposing structural restraints on three pairs of residues that are in close proximity to the pocket (Figure S1B). In addition, the excluded volume of this pocket was accounted for by placing an 8-Å-

**Received:** April 11, 2014

 © XXXX American Chemical Society  dx.doi.org/10.1021/ct500313e | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

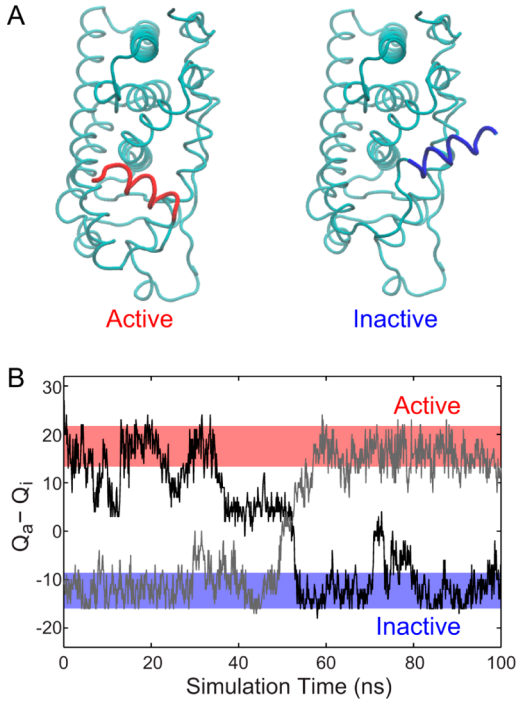



**Figure 1.** Simulating the LBD conformational switching. (A) The C-terminal $\alpha$-helix H12 switches between two distinct and well-defined conformations of the active state (in red; PDB entry 1QKU[4]) and the inactive state (in blue; PDB entry 1ERR[1]). (B) Observed transitions in simulations. The difference in contact numbers ($Q_a - Q_i$) was used to monitor the transition, where $Q_a$ is the number of contacts made in reference to the active conformation and $Q_i$ is in reference to the inactive (see Supporting Information). Plotted are two representative trajectories, which are also animated in Movie S1.

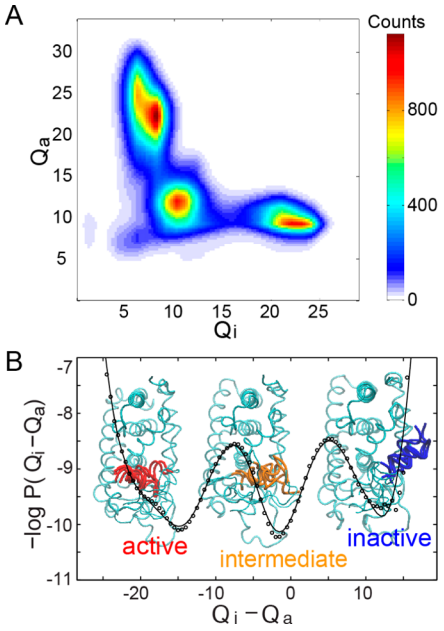



**Figure 2.** An energy landscape view of the LBD conformational transition. (A) A two-dimensional histogram plot based on the projection onto the space of $Q_i$ and $Q_a$. The upper-left corner represents an ensemble of active structures with large $Q_a$ values, while the lower-right corner represents the inactive with large $Q_i$ values. The color code of the bar below indicates the number of histogram counts from a total of 20-$\mu$s simulation trajectories. (B) A one-dimensional projection based on the histogram $P(Q_i - Q_a)$ (marked by open dots and a smoothed curve). Representative ensemble structures for the active (in red), intermediates (in orange), and the inactive (in blue).

radius sphere around the ligand-binding cavity (Figure S1C) to prevent H12 from entering the pocket (see Supporting Information). Thus, an energy model for the entire ΔLBD−linker−H12 was constructed with no prior structural information taken into account for ΔLBD−H12 interactions. As we shall see, its implementation into Langevin-based molecular dynamics (MD) simulations (see Supporting Information) is able to characterize the H12 conformational dynamics with a predictive power.

Transitions between the LBD active and inactive conformations were observed from this residue-level modeling. To keep track of the transition, a structure contact parameter of $Q_a - Q_i$ was used, where $Q_a$ and $Q_i$ are the total number of native contacts formed between ΔLBD and H12 using the active and inactive states as references, respectively (see Supporting Information). Figure 1B shows two representative trajectories, each one making a full transition between the active (indicated by red) and inactive (blue) states (see also Movie S1), thereby suggesting that this simple coarse-grained model enables a proper transition between the two conformations.

To further characterize the conformation space visited, a two-dimensional histogram of $Q_a$ and $Q_i$ was constructed based on a total of 20 $\mu$s simulation data (see Supporting Information). Figure 2A shows there are three major conformational species populated: the active state at the upper corner around $Q_a = 27$ and $Q_i = 7$, the inactive at the right area around $Q_a = 9$ and $Q_i = 23$, and the intermediate around $Q_a = 12$ and $Q_i = 11$. Similarly, this three-species distribution was also observed on a one-dimensional projection to the parameter of $Q_i - Q_a$ (Figure 2B), where three corresponding minima are readily located based on the values of $Q_i - Q_a$. There are 32 switching events between the active and inactive conformations observed in the 20-$\mu$s simulation. Note that while the convergence in both Figure 2A and B could be further examined, e.g., by running additional sets of simulation trajectories, the main focus here is to show that our energy function is able to locate the characteristic conformations. Representative ensemble structures for each state are shown in Figure 2B. Both the active (in red) and inactive (in blue) conformations are similar to corresponding crystal structures, where we observed a less than 4-Å RMSD for the H12 when ΔLBD was aligned, while a structurally diverse ensemble can be adopted within the intermediate in terms of H12 configurations (shown in orange in Figure 2B).

This predictive energy model was also applied by performing an energetic analysis on a novel D538 mutation identified recently in advanced breast cancer tumors with acquired resistance to hormonal therapy.[5,6] To assess the influence of this mutation on the conformational landscape, we selected a set of 100 random configurations for each state, where the energy change $\Delta E$ between the D538G mutant and the wild-type was calculated for each configuration (see details in the Supporting Information). Figure 3 shows this substitution of Asp538Gly yields an energy change of $\Delta E = -2.4$ kcal/mol for the active configurations (in red) and $\Delta E = -0.9$ kcal/mol for the inactive, with a net $\Delta E$ difference of $\Delta\Delta E = 1.5$ kcal/mol. We note that this energetic estimation of $\Delta E = -2.4$ kcal/mol is comparable to the energetic gain of a hydrogen-bond formation reported from atomic-level simulations at the active state.[5] These results indicate that the D538G mutation shapes


dx.doi.org/10.1021/ct500313e | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

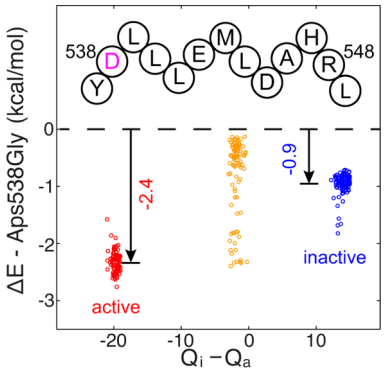



**Figure 3.** Energetic analysis of a newly identified clinical mutation D538G. The substitution stabilizes the active state (in red) by about 1.5 kcal/mol, relative to the inactive state (in blue). $\Delta E$ is the energy difference between the wild-type and the D538G mutant and in terms of $\Delta$LBD-H12 interactions. Shown on the top are the amino acids within H12, where D538 is at the beginning of H12 and highlighted in magenta. A random set of 100 configurations for each state was used for analysis; based on Figure 2A, the criterion of $25 < Q_a < 26$ and $5 < Q_i < 6$ was used to represent the active, $9 < Q_a < 10$ and $22 < Q_i < 24$ for the inactive, and $10 < Q_a < 12$ and $9 < Q_i < 11$ for the intermediate.

the entire energy landscape by stabilizing both the active and inactive conformations, although more preferably for the active. This stability readily leads to a conformation-driven activation mechanism contributing to constitutive activity of this mutant in the absence of hormones as observed in hormone-resistant breast cancer.[5] In addition, compared to previous structure modeling studies that focus on the active conformation alone,[5,6] a key finding here is the new ability of our residue-level model to explore a broad energy landscape, far beyond a single-conformation analysis.

In summary, we demonstrate that the CG model at the amino acid level is able to correctly predict the conformational switching of the functionally important LBD and that energetic calculations support the concept of a mutation-shifting landscape, which provides a conformation-driven explanation for the oncogenic D538G mutation at the detailed conformational level. While the resolution of this energetic model is at the "intermediate" level of amino acids, its ability to capture the large-scale H12 switching provides a fair validation of the underlying coarse-grained parameters used, which further adds to its test pool of complex model systems having been successfully simulated.[9] From a methodological perspective, this physics-based CG modeling adds itself to the a set of theoretical modeling methods with a predictive power.[10−12] From a conceptual level, the energetic change induced by a cancer-activating mutation is consistent with the concept of shifting landscape,[13] which provides a plausible conformation-driven mechanism for the cancer-activating D538G mutation of the ER even in the presence of therapeutic inhibitors. Above all, we demonstrate that the CG model is equipped with the ability to correctly model large-scale conformational switching to account for the inherent conformational flexibility. Ultimately, this new capability will allow us to model the interactions of this highly flexible LBD with its functional partners such as the DNA-binding domain (DBD),[14] thereby exploring the challenging territory of incorporating the inherent LBD flexibility for highly promising flexible-docking simulations occurring in the entire LBD−DBD complex.

## ■ ASSOCIATED CONTENT

### Supporting Information

Methods, one figure, one movie, one table, and additional references. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: sichun.yang@case.edu.

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Marc Parisien and Geoffrey Greene for discussion and proofreading. Computational support was provided in part by the Ohio Supercomputer Center and the CWRU high-performance computing cluster.

## ■ ABBREVIATIONS

ER, estrogen receptor; LBD, ligand-binding domain

## ■ REFERENCES

(1) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R.; Bonn, T.; Engström, O.; Ohman, L.; Greene, G. L.; Gustafsson, J.; Carlquist, M. Molecular Basis of Agonism and Antagonism in the Oestrogen Receptor Target Genes. *Nature* **1997**, *389*, 753−758.

(2) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The Structural Basis of Estrogen Receptor/coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95*, 927−937.

(3) Kumar, V.; Green, S.; Stack, G.; Berry, M.; Jin, J. R.; Chambon, P. Functional Domains of the Human Estrogen Receptor. *Cell* **1987**, *51*, 941−951.

(4) Gangloff, M.; Ruff, M.; Eiler, S.; Duclaud, S.; Wurtz, J. M.; Moras, D. Crystal Structure of a Mutant hERalpha Ligand-Binding Domain Reveals Key Structural Features for the Mechanism of Partial Agonism. *J. Biol. Chem.* **2001**, *276*, 15059−15065.

(5) Toy, W.; Shen, Y.; Won, H.; Green, B.; Sakr, R. A.; Will, M.; Li, Z.; Gala, K.; Fanning, S.; King, T. A.; Hudis, C.; Chen, D.; Taran, T.; Hortobagyi, G.; Greene, G.; Berger, M.; Baselga, J.; Chandarlapaty, S. ESR1 Ligand-Binding Domain Mutations in Hormone-Resistant Breast Cancer. *Nat. Genet.* **2013**, *45*, 1439−1445.

(6) Merenbakh-Lamin, K.; Ben-Baruch, N.; Yeheskel, A.; Dvir, A.; Soussan-Gutman, L.; Jeselsohn, R.; Yelensky, R.; Brown, M.; Miller, V. A.; Sarid, D.; Rizel, S.; Klein, B.; Rubinek, T.; Wolf, I. D538G Mutation in Estrogen Receptor-A: A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. *Cancer Res.* **2013**, *73*, 6856−6864.

(7) Robinson, D. R.; Wu, Y.-M.; Vats, P.; Su, F.; Lonigro, R. J.; Cao, X.; Kalyana-Sundaram, S.; Wang, R.; Ning, Y.; Hodges, L.; Gursky, A.; Siddiqui, J.; Tomlins, S. A.; Roychowdhury, S.; Pienta, K. J.; Kim, S. Y.; Roberts, J. S.; Rae, J. M.; Van Poznak, C. H.; Hayes, D. F.; Chugh, R.; Kunju, L. P.; Talpaz, M.; Schott, A. F.; Chinnaiyan, A. M. Activating ESR1 Mutations in Hormone-Resistant Metastatic Breast Cancer. *Nat. Genet.* **2013**, *45*, 1446−1451.

(8) Dai, S. Y.; Burris, T. P.; Dodge, J. A.; Montrose-Rafizadeh, C.; Wang, Y.; Pascal, B. D.; Chalmers, M. J.; Griffin, P. R. Unique Ligand Binding Patterns between Estrogen Receptor Alpha and Beta Revealed by Hydrogen-Deuterium Exchange. *Biochemistry* **2009**, *48*, 9668−9676.

(9) Ravikumar, K. M.; Huang, W.; Yang, S. Coarse-Grained Simulations of Protein-Protein Association: An Energy Landscape Perspective. *Biophys. J.* **2012**, *103*, 837−845.
(10) Cheng, R. R.; Morcos, F.; Levine, H.; Onuchic, J. N. Toward Rationally Redesigning Bacterial Two-Component Signaling Systems Using Coevolutionary Information. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E563−71.
(11) Zheng, W.; Schafer, N. P.; Davtyan, A.; Papoian, G. A.; Wolynes, P. G. Predictive Energy Landscapes for Protein-Protein Association. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 19244−19249.
(12) Marks, D. S.; Colwell, L. J.; Sheridan, R.; Hopf, T. A.; Pagnani, A.; Zecchina, R.; Sander, C. Protein 3D Structure Computed from Evolutionary Sequence Variation. *PLoS One* **2011**, *6*, e28766.
(13) Tsai, C.-J.; Nussinov, R. The Free Energy Landscape in Translational Science: How Can Somatic Mutations Result in Constitutive Oncogenic Activation? *Phys. Chem. Chem. Phys.* **2014**, *16*, 6332−6341.
(14) Huang, W.; Greene, G. L.; Ravikumar, K. M.; Yang, S. Cross-Talk between the Ligand- and DNA-Binding Domains of Estrogen Receptor. *Proteins* **2013**, *81*, 1900−1909.