JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

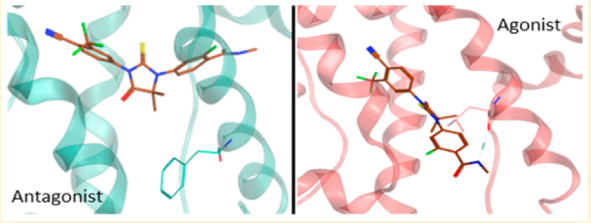
Cheminformatics Modeling of Adverse Drug Responses by Clinically Relevant Mutants of Human Androgen Receptor

Naman Paul, Lavinia A. Carabet, Nada Lallous, Takeshi Yamazaki, Martin E. Gleave, Paul S. Rennie, and Artem Cherkasov*

Vancouver Prostate Centre, Department of Urologic Sciences, Faculty of Medicine, The University of British Columbia, 2660 Oak Street, Vancouver, British Columbia, Canada V6H 3Z6

Supporting Information

ABSTRACT: The human androgen receptor (AR) is a ligand-activated transcription factor that plays a pivotal role in the development and progression of prostate cancer (PCa). Many forms of castration-resistant prostate cancer (CRPC) still rely on the AR for survival. Currently used antiandrogens face clinical limitations as drug resistance develops in patients over time since they all target the mutation-prone androgen binding site (ABS), where gain-of-function mutations eventually convert antagonists into agonists. With a significant number of reported distinct mutations located across the ABS, it is imperative to develop a prognostic platform which would equip clinicians with prior knowledge and actionable strategies if cases of previously unreported AR mutations are encountered. The goal of this study is to develop a theoretical approach that can predict such previously unreported AR mutants in response to current treatment options for PCa. The expected drug response by these mutants has been modeled using cheminformatics methodology. The corresponding QSAR pipeline has been created, which extracts key protein−ligand interactions and quantifies them by 4D molecular descriptors. The developed models reported with an accuracy reaching 90% and enable prediction of activation of AR mutants by its native ligand as well as assess whether known antiandrogens will act on them as agonists or antagonists. As a result, a previously uncharacterized mutant, T878G, has been predicted to be activated by the latest antiandrogen enzalutamide, and the corresponding experimental evaluation confirmed this prediction. Overall, the developed cheminformatics pipeline provides useful insights toward understanding the changing genomic landscape of advanced PCa.



INTRODUCTION

The androgen receptor (AR) is activated by binding of androgen steroids, such as 5α- dihydrotestosterone (DHT) and testosterone, to the androgen binding site (ABS) of the receptor that results in its nuclear translocation and thus its DNA binding and transcriptional activation. The ABS serves as the primary target of antiandrogens, such as bicalutamide, hydroxyflutamide, and enzalutamide, administered post androgen deprivation therapy (ADT).[1−3] A number of mechanisms for drug resistance have been investigated, including modifications in AR cofactor balance, elevated AR protein levels, androgen synthesis by tumor cells, expression of truncated AR variants and point mutations.[4,5] Many mutations were ascribed to drug resistance, which develops in patients over time with the use of antiandrogens and eventually, converts the drug response from antagonist to agonist.[6] This has been observed in first-generation as well as second-generation antiandrogens such as enzalutamide.[7,8] Thus, the importance of identification and functional characterization of AR mutants cannot be ignored, making it an essential component in predicting and monitoring the therapeutic response of patients.[9] These AR mutations are located across the ligand binding domain (LBD) in castration resistant prostate cancer (CRPC) patients and respond distinctly to different antiandrogens. The in vitro characterization of reported AR mutants located across the LBD, identified by cell-free DNA (cfDNA) sequencing, revealed new gain-of-function scenarios.[10,11] Therefore, it is vital to further investigate AR mutations and identify new ones responsible for antiandrogen treatment failure in order to build up a prognostic platform using evidence-based approaches.

This study utilizes a synergetic combination of modern computer-aided drug discovery (CADD) tools, cheminformatics and machine learning approaches to model AR mutants, in terms of their transcriptional activation and their therapeutic response to the currently used antiandrogens.[12−16] Initially, the AR mutants were engineered in-silico, followed by molecular docking of clinically used antiandrogens and the native ligand DHT into the ABS, to capture their binding conformations (as shown in Figure 1).

The resulting protein−ligand interactions were then scored using recently developed 4D inductive descriptors. Through the established exploratory data analysis pipeline (see Figure 2), we constructed quantitative structure activity relationship (QSAR) models, to predict the ligand based activation of AR mutants and to classify the drug response as agonist or antagonist.[17] The molecular dynamics simulations were further employed to analyze the structural integrity of predicted mutants and

Received: July 11, 2016
Published: November 7, 2016

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

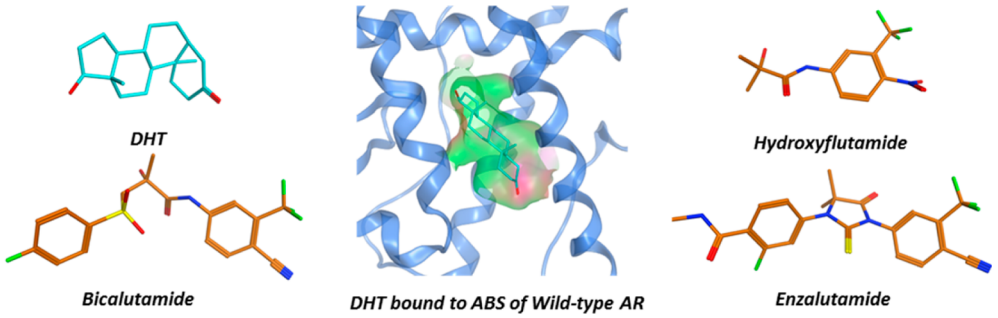


Figure 1. Native ligand DHT (cyan) bound to the androgen binding site (ABS) of wild type AR, which serves as the target site for other shown antiandrogens.

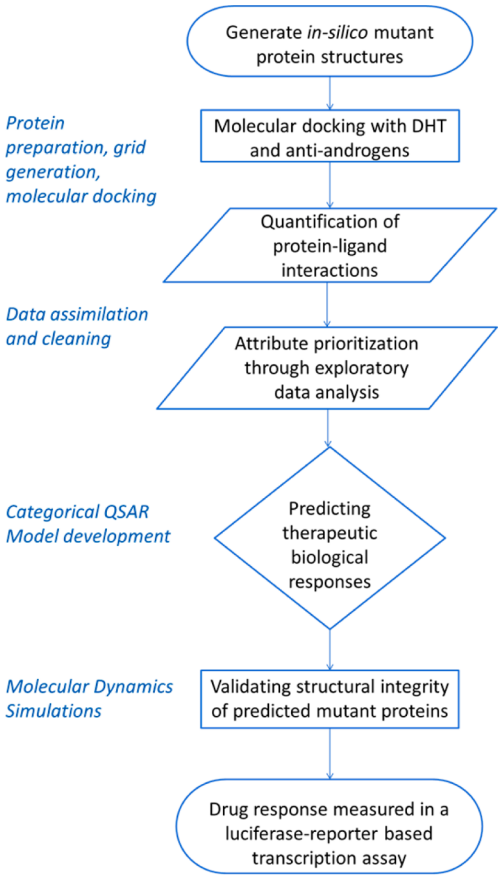


Figure 2. Pipeline developed for predicting activity of AR mutants and their drug responses.

mutant-drug complexes.[18] Following substantial in-silico analyses and validation, to further verify our predictions, AR mutants were experimentally created and tested in cell-based assays for their transcriptional activation and drug response. This study aims at predicting drug response profiles of AR mutants not formerly reported but might occur in patients undergoing antiandrogen therapies. In this work, we report a previously uncharacterized mutant T878G located in AR-LBD, which is activated by both first as well as second generation antiandrogens used in the current clinical practice. These antiandrogens exhibit agonist responses against the mutant as both predicted and experimentally validated. These results indicate that the developed approach should provide insights for clinicians to look out for such novel mutants, and to administer drugs that would yield an anticipated response in a patient-specific manner.

MATERIALS AND METHODS

Mutant Data Set Generation and Modeling. A cheminformatics pipeline was developed for carrying out this study, as illustrated by Figure 2. A protein model for the wild-type (WT) AR was built based on PDB ID: 1Z95 crystallographic structure, subjected to addition of missing hydrogen atoms and energy minimization.[6] The resulting energy equilibrated wild-type structure was used as a template to generate mutant protein structures using molecular operating environment (MOE) software.[19] Functional characterization experiments were recently carried out at the Vancouver Prostate Centre to understand the response of PCa drugs on 24 known AR mutants and the wild-type AR. The corresponding responses were categorically classified to be either agonist (+1) or antagonist (−1). These responses were assigned to the modeled protein structures complexed with various drug molecules via molecular docking and subsequently used to create the 4D QSAR training set.[11] (See Table S1) In order to investigate novel mutants within the ABS, with unavailability of any prior knowledge about their stability and characteristic response to antiandrogens a total of 28 amino-acid residues located within the ABS were mutated by 19 other side chain substitutions. A total of 532 single point mutant AR structures were generated by this approach. These structures were then docked with antiandrogens and protein−ligand interactions were quantified, making up the test set, for which response predictions were generated through QSAR analysis.

The protein structures were prepared using Protein Preparation Wizard module of Schrödinger's Maestro suite. Water molecules were removed from the crystal structure, followed by hydrogen bond assignment, and subsequent restrained energy minimization using OPLS_2005 force field.[20] Eventually 22 of these structures were eliminated, after protein preparation since certain substitutions cause steric clashes and structural disruption of proteins. Receptor grids were generated corresponding to the crystallographic drug molecule present within protein structure. The van der Waals radius scaling factor was reduced to 0.8 in order to soften the potential for nonpolar parts of the receptor. Glide XP (extra precision) docking protocol was applied for molecular docking.[21−23] The Glide per-residue-scores were also calculated for residues lying within 10 Å of the ligand, and used as molecular descriptors. This procedure was implemented for both training and the test sets. In addition to the native ligand DHT, antiandrogens used for molecular docking included: ARN-509, bicalutamide, enzalutamide, and hydoxyflutamide.[24−27]

Novel 4D-Inductive Descriptors. Molecular descriptors are the central piece of QSAR modeling and have been used to describe chemical features of molecules, over various levels of structural representation.[28] Two- and three-dimensional descrip-

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

tors have been widely applied to diverse QSAR modeling pipelines that predict biological activities and binding affinities of new ligand molecules among other end points.[29] However, these ligand descriptors do not quantify the effect of ligand on the receptor or vice versa. On the basis of our previous three-dimensional (3D) models of inductive effects (σ), steric effects (Rs, Abs_Rs), inductive electronegativity (χ), inductive charge (Q), and molecular capacitance, a range of four-dimensional (4D) inductive descriptors has been employed for this study to quantify receptor−ligand interactions at the atomic level.[30] The theory and mathematical formalism of the inductive descriptors has been described elsewhere, and they have been successful in conveying required ligand's chemical effect description.[31−35] All the inductive descriptors possess obvious physical meaning and can be readily calculated from fundamental and accessible atomic properties such as covalent radii, electronegativity, and interatomic distances, making the inductive descriptors particularly useful for QSAR studies in large sets of chemical compounds and sizable molecular systems such as proteins as well as their complexes.

The 4D inductive descriptors were implemented in the scientific vector language (SVL)[36] and computed by our in-house script integrated in MOE. Three categories of descriptors were developed and calculated for all receptor−ligand complexes: (1) Rs_L_R, Abs_Rs_L_R, σ_L_R, and Q_L_R descriptors were calculated to measure, the steric, polar, and induced charge respectively, the cumulative influence of all ligand (L) atoms on all receptor (R) atoms within 10 Å receptor region surrounding the ligand; (2) AA#_Rs_AA_R, AA#_Abs_Rs_AA_R, and AA#_σ_AA_R descriptors were computed to quantify the steric and inductive effects of all atoms of mutated amino acid residues (AA#) relative to the wild-type; and (3) R_Ah_L_Ah descriptors were calculated to quantify, as the inverse square of interatomic distances, the interactions between all receptor atoms within the 10 Å cutoff region in all possible hybridized states (R_Ah) and all ligand atoms in their hybridized states (L_Ah).

The significance and interpretation of our novel 4D inductive descriptors in discriminating the behavior of AR mutants relative to the AR wild-type are highlighted in the Results and Discussion section.

Descriptor Pruning and Data Cleaning. The prioritization and ranking of molecular descriptors (attributes) for model building was carried out using a new exploratory data analysis pipeline developed in the R programming language.[37,38] A total of 225 attributes were analyzed for their importance in order to establish the optimum number of attributes that can be utilized for categorical classification. The attributes were sorted for their relevance by numerical estimates of importance based on random forest classification algorithm and then prioritized using Boruta package. The optimum number of attributes was then determined through recursive feature elimination implemented through the caret package.[39−41] The downsizing of the number of descriptors, accounts for the removal of highly correlated descriptors and creation of a better attribute pool. This prioritization approach maximizes the statistical contribution of attributes that enhances the QSAR modeling efficiency.

QSAR Modeling Pipeline. QSAR modeling technique has been very successful in making biologically relevant predictions, contributing to discovery of new inhibitors for a wide range of targets.[42] Generally, continuous QSAR models predict biological activity of novel compounds based on prior knowledge of biological activities of already studied molecules. With our goal being classification of drug responses, two types of binary categorical QSAR models were deployed: (1) for predicting protein activation−inactivation upon DHT stimulation and (2) for predicting the drug response with mutant proteins to either an agonist or antagonist response. The following seven classification algorithms implemented in WEKA, including LibSVM, IBk, Bagging, Dagging, RandomForest, OneR, and DecisionStump were used to construct the corresponding QSAR models.[39,43−49] Importantly, a consensus vote approach was utilized to examine instance-predictions and to reduce chances of biased predictions by any particular algorithm. These models were evaluated for their accuracy by 10-fold cross validation as well as external test set validation.

Applicability Domain Assessment. Applicability domain assessment was then employed to assess QSAR model performance. The uniformity in distribution of molecular descriptor values was evaluated between both the training and test set instances. Utilizing this information, the exactness of predictions can be estimated for any random data set with minimal similarity consideration used for model validation.[50] The descriptor space for both train and test sets was evaluated between the minimum and maximum descriptor values (±15%). The presumption of this procedure is that predictive accuracy for those instances that lie within the range of descriptor values (±15%) tend to be higher compared to those that are beyond the permissible threshold (see Figure S3).

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were carried out using Amber force field system, installed on Westgrid High Performance Computing Clusters in Canada, to evaluate the structural integrity of mutant receptor−ligand complexes and to verify the contacts between the receptor and ligand obtained from molecular docking. The following steps describe the MD workflow:[51]

1. The geometry optimization was carried out for the ligand and the electrostatic potential around the optimized ligand structure was calculated with Gaussian09 program available on Westgrid's Grex cluster.[52]
2. Antechamber was used to assign atomic charges based on RESP fitting and GAFF atom types to the ligand molecule.[53−55]
3. The TLEAP program was employed for assignment of protein−ligand force fields, to neutralize the system by adding counterions depending upon the total charge of the system, and to solvate the system in a TIP3P 10.0 Å water box.[56]
4. The system was then energy minimized to remove steric clashes, if any, in a four-step process, with sequentially decreasing restraint weights from 50 to 0 kcal/mol-Å^2.
5. The system was then heated from an initial temperature of 100 up to 300 K with a fixed protein structure.
6. After heating, the protein structure was still fixed and subjected to gradual reduction of protein constraints in three-step process, while a pressure measuring 1 ATM was applied to the system for density equilibration.
7. Following equilibration, a 25 ns production run was initiated. All of the above heating, density equilibration, and production runs were performed on 24 processors across 2 nodes on Westgrid's Jasper cluster with a wall time of 72 h and 2000 MB memory allocation. The production run finished in 40 h on an average, for the mutant protein complexes (see the Supporting Information for details).

In-Vitro Screening. The AR mutants predicted by the QSAR models were experimentally created and their response to increasing concentrations of various antiandrogens has been

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

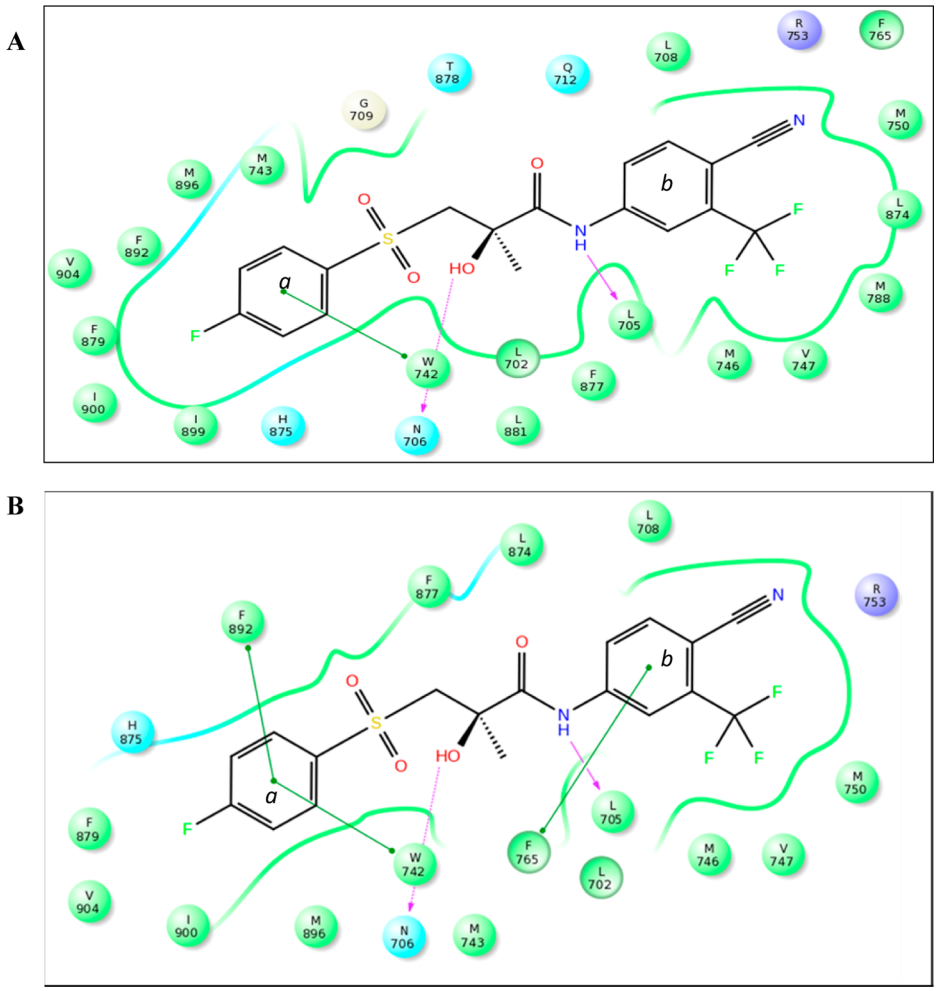


Figure 3. (A) Bicalutamide interactions with wild type AR, involving L705, N706, and W742 residues. (B) Additional interactions seen in T878G mutant, especially the $\pi-\pi$ stacking of benzene rings of residues F892 and F765 with a and b rings of bicalutamide.

measured using a luciferase-reporter transcriptional assay.[11] PC3 cells lacking the AR and were maintained in RPMI 1640 media (Life Technologies) and 5% FBS (Hyclone ThermoFisher Scientific) at 37 °C and 5% CO_2. For the AR inhibition assay, cells were seeded in 96-well plates (5,000 cells/well) in RPMI 1640 medium with 5% charcoal-stripped serum (CSS) (Hyclone). Twenty-four hours after seeding, cells were cotransfected with 25 ng of wild-type or mutated AR and 25 ng of the reporter plasmid pARR3-tk-luciferase that is under the transcriptional control of AR. After 48 h, cells were stimulated with 0.1 nM of the synthetic androgen, R1881, and treated with 0.1% DMSO (for the control) or increasing concentrations of antiandrogens (bicalutamide, enzalutamide, and hydroxyflutamide) ranging from 0 to 50 μM. The medium was aspirated after 24 h and cells were lysed by adding 60 μL of 1× passive lysis buffer (Promega). The luciferase activity was quantified as a luminescence signal measured by a TECAN M200Pro instrument, after adding 50 μL of luciferase assay reagent (Promega) to 20 μL of lysate. Each concentration was assayed in quadruplicate, with 2−3 biological replicates. Results were normalized to the wild-type AR activity.

RESULTS AND DISCUSSION

Molecular Docking Reveals Distinct Binding Poses. A test set comprising of 532 mutant structures was constructed and 22 structures were removed from the test set in the protein preparation step. Consequently, receptor grids for the remaining 510 mutant structures and 25 training set structures (535 in total) were generated. Ligand molecules were docked into the binding site, using Glide XP (extra precision) mode enabling the precise pose prediction and hence the elimination of false positives in terms of docked poses. Compared to Glide SP (standard precision) mode, a higher amount of computational resources were consumed. A total of 2140 docked structures were exported along with their Glide per-residue-scores, for the external test set and 84 for the training data set

Several van der Waals and hydrogen-bond interactions were observed between the antiandrogen bicalutamide and wild-type AR (see Figure 3A). The hydroxyl group of bicalutamide interacts with side chain of N706, whereas the backbone of L705 exhibited H-bonding with propanamide moiety. These observations are consistent with those previously reported by Bohl et al., 2005[6] while studying bicalutamide resistance and give confidence in the quality and accuracy of the obtained docking poses. Importantly, the mutant protein−ligand complexes demonstrate altered interactions with antiandrogens in contrast to the wild-type AR that may correlate with the agonist feedback.

A previously uncharacterized mutant T878G predicted in this study showcases the aforementioned altered interactions. The substitution of threonine at the 878th position into glycine increases the ligand accessible surface area in the ABS pocket, thus allowing elevated levels of ligand−receptor interaction.

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

Table 1. Variance in 4D-Inductive Descriptor Values That Contributes to Behavioral Drug Response Trends[a]

H875Y	Rs_L_R	Abs_Rs_L_R	875 Rs_AA_L	R_Csp3-L_Fsp3	R_Nsp2-L-Nsp
bicalutamide	239.39	70.57	0.37	13.37	1.17
enzalutamide	268.38	78.94	0.48	13.12	1.16
hydroxyflutamide	157.73	46.65	0.19	9.84	0.00

[a]Rs_L_R: steric hindrance on the receptor (within 10 Å) caused by the ligand. Abs_L_R: absolute steric hindrance on the receptor (within 10 Å) caused by the ligand. 875 Rs_AA_L: steric hindrance on residue 875 of the receptor caused by the ligand. R_Csp3-L_Fsp3: steric hindrance on sp3 hybridized carbons of the receptor caused by sp3 hybridized fluorines of the ligand. R_Nsp2-L_Nsp: steric hindrance on sp2 hybridized nitrogens of the receptor caused by sp hybridized nitrogens of the ligand.

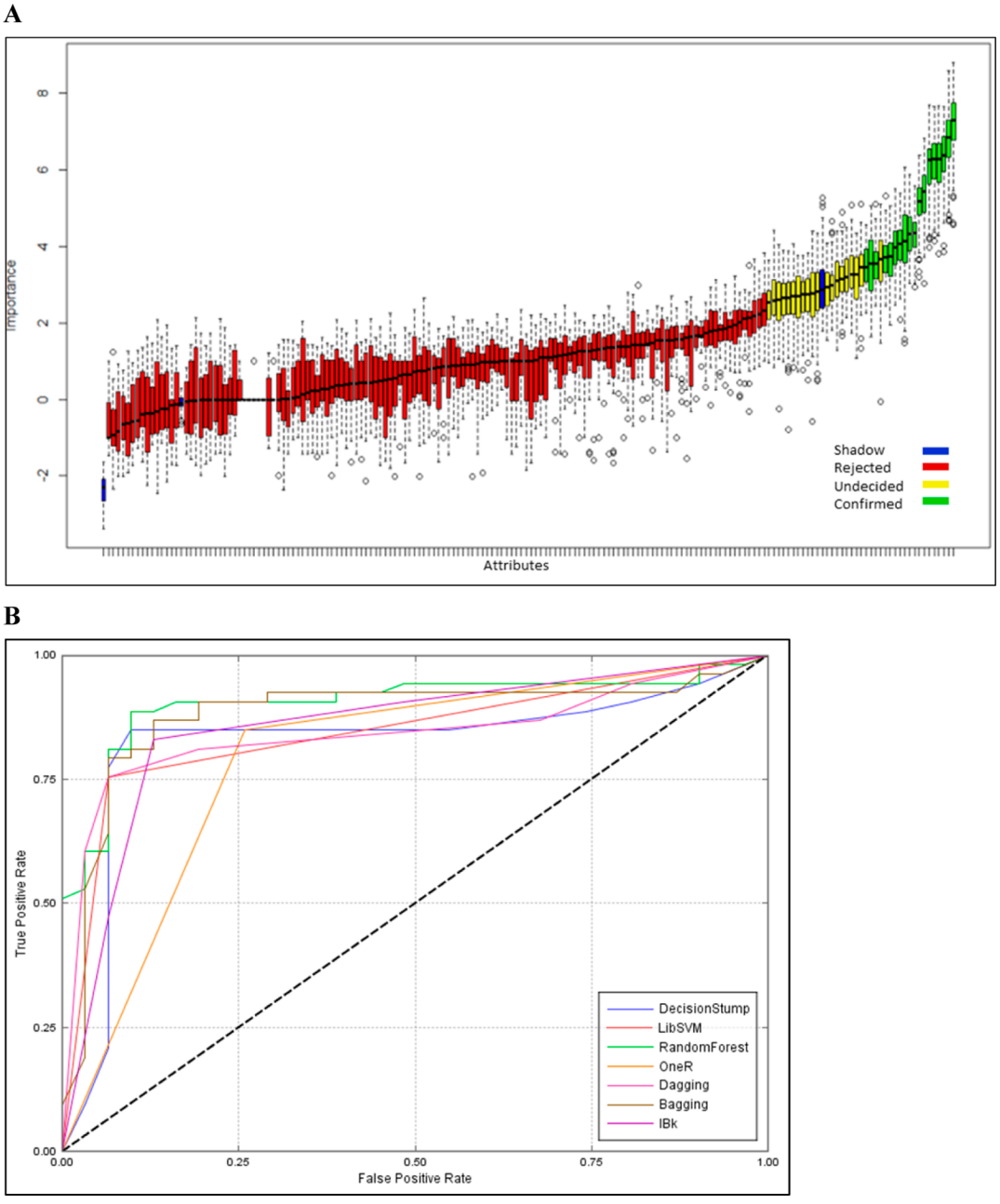


Figure 4. (A) Prioritization of attributes, implemented through Boruta package. Blue boxplots depict minimum, average, and maximum Z score of a shadow attribute. Red and green boxplots show rejected and confirmed attributes respectively whereas, the yellow boxplots represent the undecided class. (B) Over 85% area under the receiver operator characteristic (ROC) curve shows the high sensitivity and specificity of the classifiers.

In addition to the interactions in the wild-type−bicalutamide complex, in the T878G mutant, the benzene rings of the F892 and F765 undergo $\pi-\pi$ stacking with both *a* and *b* rings of bicalutamide (see Figure 3B). Analogous interactions of this mutant with enzalutamide is discussed in subsequent sections (also see Figure S1).

Quantification of Receptor−Ligand Interactions Reveals Statistical Trends. To thoroughly examine the wild-type−mutant

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

interaction divergence, the receptor−ligand contacts were quantified within 10 Å range using 4D-inductive descriptors. These molecular descriptors were merged with the Glide per-residue-scores for all complexes, yielding a total of 225 QSAR parameters.

The 4D-inductive molecular descriptor "Rs_L_R" represents the steric hindrance faced by ligand molecules bound to a specific mutant receptor. For example, AR mutant H875Y presents a strong correlation between biological responses and descriptor values. It has been reported that hydroxyflutamide yields strong agonist response toward this mutant, whereas, for larger ligands like bicalutamide and enzalutamide, the agonist response is relatively weaker.[11,57] This correlation of attribute values with antiandrogen response trends of the mutant H875Y can be seen in Table 1. This variance in response can be attributed to the amount of steric hindrance faced by each of these antiandrogens, the least being faced by hydroxyflutamide that correlates with its feedback response.

The next essential step was the prioritization of these attributes, which was achieved by importance ranking and pruning. The removal of highly correlated, redundant, statistically insignificant attributes enhances prediction efficiency and decreases the computational complexity.[58]

Through the Boruta package, three shadow attributes were added to the data set, (shown in blue, see Figure 4A) and Z-scores were accumulated through random forest algorithm implementation. In the initial run, the attributes were then segregated into three categories of importance (confirmed, undecided, and rejected). 182 attributes with lower importance than MZSA (maximum Z-score shadow attribute) were filtered in the first run whereas the "undecided" class was re-evaluated through 99 iterations across 22.70 s. Ultimately 28 attributes were confirmed to be important and 197 were removed after the second run. The recursive feature elimination (RFE) method of the caret package was used to determine the number of attributes required for efficient model building based on lowest RMSE (root-mean-square error) measured across different number of attributes. Therefore, the following four attributes along with the glide e-model score were selected for QSAR model building: R_Csp3-L_Nsp3, R_Csp2-L_Nsp3, R_Osp3-L_Nsp3, and R_Osp2-L_Nsp3 (see Table S2 and Figure S2 for further details).

Furthermore, applicability domain assessment of these molecular descriptors determined them to be well within the chemical space of the minimum−maximum range threshold values of ±15% (see Figure S3).

QSAR Models Predict Drug Response of Mutants. QSAR models built for categorical classification, predict the activity/inactivity of the mutants in response to androgen steroids, and the agonist/antagonist response to antiandrogen drugs. A nominal attribute "active" was employed for modeling active (1) or inactive (0) responses upon DHT stimulation. The nominal attribute of "activity" for agonist/antagonist QSAR model was assigned with either agonist (+1) or antagonist (−1) values obtained from precharacterized AR mutants response to antiandrogens. This classification was used as the training data set for the agonist/antagonist effect model.

A 10-fold cross validation approach was employed to validate the agonist/antagonist QSAR model predictions. Table 2 below features the statistical performance for all algorithms employed in QSAR model building. The accuracy value was over 82% for all the classifiers, with highest being 88.09% for LibSVM classifier algorithm. The calculated average sensitivity was ∼84%, which reflects the model efficiency in true positive identification.

Table 2. Performance Statistics upon 10-fold Cross Validation of the 4D QSAR Training Data Set Employing Classification Algorithms Used for Model Building

s. no.	algorithms	sensitivity	specificity	PPV	accuracy	ROC
1	Bagging	0.86	0.87	0.90	86.90%	0.89
2	Dagging	0.75	0.93	0.95	82.14%	0.89
3	DecisionStump	0.90	0.90	0.93	86.90%	0.82
4	IBk	0.79	0.87	0.91	82.14%	0.85
5	LibSVM	0.86	0.90	0.93	88.09%	0.88
6	OneR	0.86	0.83	0.90	85.71%	0.85
7	RandomForest	0.86	0.87	0.92	86.90%	0.87

For each classifier, the receiver operator characteristic (ROC) curve was generated to determine the accuracy in distinguishing between true and false positives. A perfect ROC curve would cover an area under the curve of 100% while below 50% coverage (as shown by dashed line in Figure 4B) denotes random outcomes which are statistically insignificant. In our case, the ROC curve, covers on an average, an area of 86.42% for all 7 classifiers. The major advantage that can be interpreted from the ROC characteristic curve is that the diagnostic power of these classifiers utilized for QSAR modeling lie above the "chance" level that adds more confidence to the predicted outcomes. The classifiers precisely identify true and false positives illustrative of the discriminant capabilities of the classifiers.

T878G Mutant Yields Agonist Response to Antiandrogens. The novelty of this study is the non retrospective QSAR analysis of mutants, without any prior knowledge about the physical feasibility or functional characterization of mutants. The QSAR models predicted a total of 12 mutants that should yield agonist response to antiandrogens. Thus, these mutants were experimentally created and experimentally tested. About nine of these mutants were biologically inactive upon testing their response to DHT and antiandrogens. Also, two mutants were weakly stimulated by DHT and with low strength signals with respect to antiandrogens, hence could not be further investigated.

We report here a novel AR-LBD mutant T878G, which QSAR models predicted to be active, upon DHT stimulation and to yield an agonist response toward currently used antiandrogens. As discussed in the previous section, the computational modeling revealed dissimilar mutant-ligand binding interactions compared to the wild-type. The T878G−enzalutamide complex was no exception to our finding. The G878 backbone makes hydrogen bond with Enzalutamide's benzamide motif. The substitution of wild-type threonine by glycine exposes the amino acid backbone that directly interacts with the ligand molecule and could potentially explain the agonist activity predicted upon enzalutamide treatment. The $\pi-\pi$ stacking interactions of F892 and F765 benzene rings with both "*a*" and "*b*" rings of enzalutamide (see Figure 5), in T-shaped conformations represent the energy minima of the protein−ligand binding geometry.

Although the magnitude of a single $\pi-\pi$ interaction is not very large, collectively these interactions may significantly influence the tertiary structure of T878G mutant.[59] By the quantification of these interactions and the consequential analysis for the T878G mutant, a correlation with the predicted therapeutic responses could be ascertained (see Table 3). This mutant was predicted to transform hydroxyflutamide into a complete agonist, with least

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

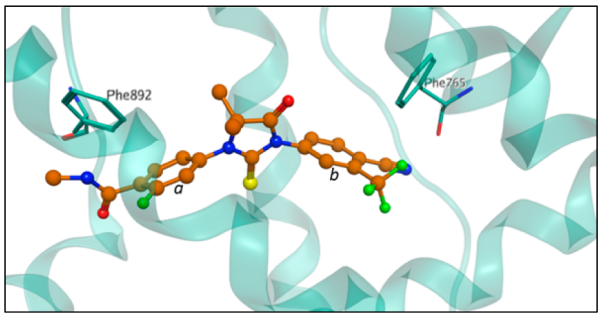


Figure 5. T878G−enzalutamide complex. $\pi-\pi$ stacking interactions can be seen between the benzene rings of F892 and F765. They are positioned in a T-shaped orientation over the aromatic ring moieties "*a*" and "*b*" of enzalutamide.

steric hindrance value of 139.23 within the ABS compared to other antiandrogens (see Figure S1). Molecular dynamics simulations were used to examine the predicted protein−ligand interactions by taking into account of the conformational flexibility and to elucidate the structural equilibrium of the complex. The binding poses of antiandrogen−T878G complexes predicted by Glide XP program were selected as a starting point for 25 ns production runs. The structural stability of T878G−enzalutamide complex was assessed by calculating the RMSD of the receptor throughout the MD simulation using the initial pose as the reference structure. An average RMSD measure of 1.1 Å over 25 ns simulation time for the average distance between the backbone atoms of superposed protein structures is reflective of the overall structural stability and integrity of the structure (see Figure 6A). Similar observations were made in cases of bicalutamide and hydroxyflutamide complex stability (see Figure S4).

An average RMSD value of 1.9 Å for enzalutamide demonstrates its binding fitness within the mutant receptor ABS, maintaining consistent orientation as with the initial conformation. In order to see how the protein flexibility adjusts the ligand docking pose, we evaluated the frequency of contact between the protein and the ligand during the 25 ns production run, and compared it with the initial docking pose. Here, we defined the contact between the ligand and a residue when the distance between any of ligand atoms and residue atoms is within 3 Å (Figure 6B). A contact frequency as high as 99.80% was obtained for F765 residue. Within helix 11 of the T878G mutant structure, the following residues presented highly consistent interactions with enzalutamide, between both initial and final points of evaluation: H875 (97.44%), F877 (91.56%), G878 (99.32%), and F892 (97.72%). Importantly, those interactions were not observed in the wild-type complexes. Therefore, these interactions are indicative of the anticipated antagonist to agonist switch in T878G mutant−antiandrogen response.

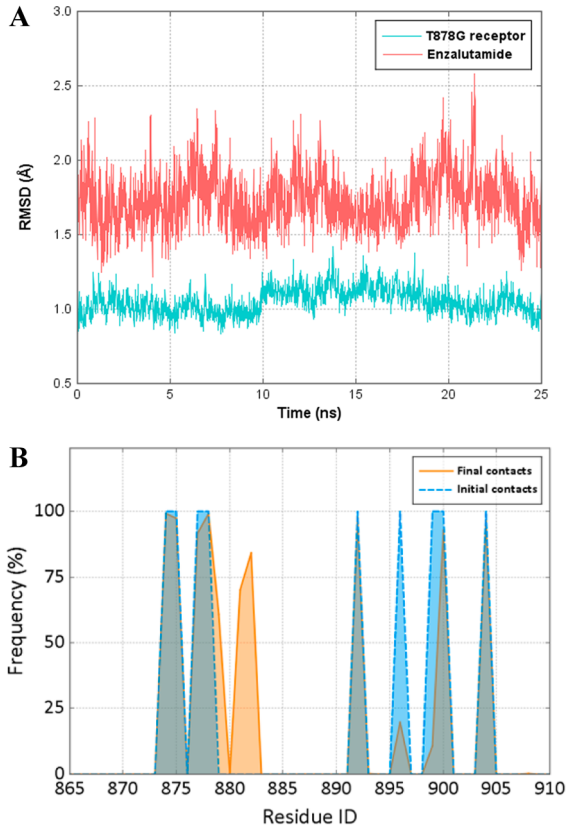


Figure 6. (A) Structural stability of T878G−enzalutamide complex, calculated through RMSD over 25 ns MD simulations. (B) Frequency of contacts is retained, between both initial docked conformation (blue) of receptor−ligand complex and MD equilibrated structure (orange). Overlapping contacts are shown in gray.

In-Vitro Screening of T878G Mutant Confirms the Predicted Drug Responses. In order to validate the in silico predictions, the response of T878G mutant to antiandrogens was evaluated experimentally. Therefore, the mutant was created using site-directed mutagenesis approach and WT-AR as the template. The response of T878G mutant to increasing concentrations of antiandrogens was evaluated in PC3 human prostate cancer cells lacking the endogenous AR. Cells were transiently transfected with either wild-type or T878G mutated AR and a luciferase reporter plasmid driven by the ARR_3-tk probasin-based promoter. Cells were stimulated with the nonmetabolizable androgen R1881 and then treated with the increasing concentrations of enzalutamide, hydroxyflutamide, or bicalutamide. As predicted T878G mutant showed an agonist response in the presence of the 3 evaluated drugs (see Figure 7).

Table 3. Descriptor Values Computed for T878G Mutant[a]

T878G	Rs_L_R	Abs_Rs_L_R	878 Rs_AA_L	R_Csp3-L_Fsp3	R_Nsp2-L-Nsp
bicalutamide	208.26	61.83	0.39	11.86	1.07
enzalutamide	231.75	68.67	0.65	11.76	1.03
hydroxyflutamide	139.23	41.47	0.21	8.93	0.00

[a]These reflect drug response trends, that were validated through QSAR modeling and in-vitro functionalization experiments. Rs_L_R: steric hindrance on the receptor (within 10 Å) caused by the ligand. Abs_L_R: absolute steric hindrance on the receptor (within 10 Å) caused by the ligand. 878 Rs_AA_L: steric hindrance on residue 878 of the receptor caused by the ligand. R_Csp3-L_Fsp3: steric hindrance on sp3 hybridized carbons of the receptor caused by sp3 hybridized fluorines of the ligand. R_Nsp2-L_Nsp: steric hindrance on sp2 hybridized nitrogens of the receptor caused by sp hybridized nitrogens of the ligand.

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

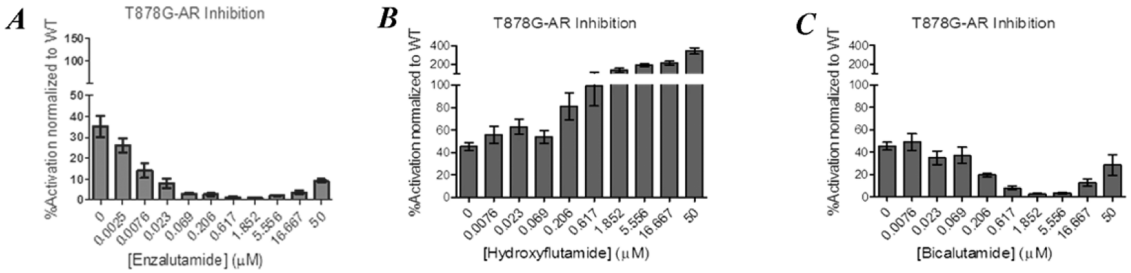


Figure 7. Response of T878G mutant to enzalutamide, hydroxyflutamide, and bicalutamide, in an in-vitro cell-based assay. Each concentration was assayed in quadruplicate $n = 4$, with a biological replicate of $n = 2$. Results were averaged and normalized by expressing them as a percentage of WT AR activity ± SEM.

CONCLUSIONS

In this study, a 4D QSAR model was developed for the in-silico functional characterization of AR mutants, and their adverse drug responses. The current methodology accurately describes and quantifies the important structural aspects involved in the mutant−drug interactions. Through our QSAR pipeline combined with biological testing, we have successfully predicted and validated the drug responses of a previously uncharacterized human androgen receptor mutant T878G. The drug resistance through agonist-like behavior exhibited by the T878G mutant reveals the importance of monitoring patients for such mutations. This approach will help predict similar undiscovered and uncharacterized mutants thereby helping in the development of alternative treatment strategies to either delay or overcome drug resistance, in clinical practice.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00400.

Additional tables containing detailed information on the selected descriptors and QSAR data set. Additional figures depicting the ligand interaction diagrams for enzalutamide and hydroxyflutamide with T878G mutant protein, applicability domain assessment, and the RMSD analysis for structural stability of the mutant with bicalutamide and hydroxyflutamide (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel.: 604-875-4111 x69628. Fax: 604-875-5654. E-mail: artc@interchange.ubc.ca.

Author Contributions

The in vitro screening experiments were carried out by N.L. The SVL script for 4-D Inductive descriptors was created by L.C. All authors have given approval to the final version of the manuscript.

Funding

The authors acknowledge financial support from the Canadian Cancer Society Research Institute (Grant no. 701585). N.L. acknowledges the support by the National Cancer Institute, SPORE program (Grant no. P50CA097186).

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

3D, three-dimensional; 4D, four-dimensional; ABS, androgen binding site; ADT, androgen deprivation therapy; AR, androgen receptor; AUC, area under the curve; CADD, computer-aided drug discovery; CRPC, castration-resistant prostate cancer; CSS, charcoal stripped serum; DHT, 5α-dihydrotestosterone; DMSO, dimethyl sulfoxide; FBS, fetal bovine serum; GAFF, general amber force fields; LBD, ligand binding domain; MD, molecular dynamics; MOE, molecular operating environment; MZSA, maximum Z-score shadow attribute; PCa, prostate cancer; PDB, protein data bank; QSAR, quantitative structure activity relationship; RESP, restrained electrostatic potential; RMSD, root-mean-square deviation; ROC, receiver operating characteristic; SEM, standard error of the mean; WT, wild type; XP, extra precision

REFERENCES

(1) Gao, W.; Bohl, C. E.; Dalton, J. T. Chemistry and Structural Biology of Androgen Receptor. *Chem. Rev.* **2005**, *105* (9), 3352−70.

(2) Wirth, M. P.; Hakenberg, O. W.; Froehner, M. Antiandrogens in the Treatment of Prostate Cancer. *Eur. Urol.* **2007**, *51* (2), 306−13 discussion 314.

(3) Hoffman-Censits, J.; Kelly, W. K. Enzalutamide: A Novel Antiandrogen for Patients with Castrate-Resistant Prostate Cancer. *Clin. Cancer Res.* **2013**, *19* (6), 1335−9.

(4) Zong, Y.; Goldstein, A. S. Adaptation or Selection–Mechanisms of Castration-Resistant Prostate Cancer. *Nat. Rev. Urol.* **2012**, *10* (2), 90−8.

(5) Wyatt, A. W.; Gleave, M. E. Targeting the Adaptive Molecular Landscape of Castration-Resistant Prostate Cancer. *EMBO Mol. Med.* **2015**, *7* (7), 878−94.

(6) Bohl, C. E.; Gao, W. Q.; Miller, D. D.; Bell, C. E.; Dalton, J. T. Structural Basis for Antagonism and Resistance of Bicalutamide in Prostate Cancer. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (17), 6201−6206.

(7) Korpal, M.; Korn, J. M.; Gao, X. L.; Rakiec, D. P.; Ruddy, D. A.; Doshi, S.; Yuan, J.; Kovats, S. G.; Kim, S.; Cooke, V. G.; Monahan, J. E.; Stegmeier, F.; Roberts, T. M.; Sellers, W. R.; Zhou, W. L.; Zhu, P. An F876l Mutation in Androgen Receptor Confers Genetic and Phenotypic Resistance to Mdv3100 (Enzalutamide). *Cancer Discovery* **2013**, *3* (9), 1030−1043.

(8) van Soest, R. J.; van Royen, M. E.; de Morree, E. S.; Moll, J. M.; Teubel, W.; Wiemer, E. A.; Mathijssen, R. H.; de Wit, R.; van Weerden, W. M. Cross-Resistance between Taxanes and New Hormonal Agents Abiraterone and Enzalutamide May Affect Drug Sequence Choices in Metastatic Castration-Resistant Prostate Cancer. *Eur. J. Cancer* **2013**, *49* (18), 3821−30.

(9) Robinson, D.; Van Allen, E. M.; Wu, Y. M.; Schultz, N.; Lonigro, R. J.; Mosquera, J. M.; Montgomery, B.; Taplin, M. E.; Pritchard, C. C.; Attard, G.; Beltran, H.; Abida, W.; Bradley, R. K.; Vinson, J.; Cao, X.; Vats, P.; Kunju, L. P.; Hussain, M.; Feng, F. Y.; Tomlins, S. A.; Cooney, K. A.; Smith, D. C.; Brennan, C.; Siddiqui, J.; Mehra, R.; Chen, Y.;

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

Rathkopf, D. E.; Morris, M. J.; Solomon, S. B.; Durack, J. C.; Reuter, V. E.; Gopalan, A.; Gao, J.; Loda, M.; Lis, R. T.; Bowden, M.; Balk, S. P.; Gaviola, G.; Sougnez, C.; Gupta, M.; Yu, E. Y.; Mostaghel, E. A.; Cheng, H. H.; Mulcahy, H.; True, L. D.; Plymate, S. R.; Dvinge, H.; Ferraldeschi, R.; Flohr, P.; Miranda, S.; Zafeiriou, Z.; Tunariu, N.; Mateo, J.; Perez-Lopez, R.; Demichelis, F.; Robinson, B. D.; Schiffman, M.; Nanus, D. M.; Tagawa, S. T.; Sigaras, A.; Eng, K. W.; Elemento, O.; Sboner, A.; Heath, E. I.; Scher, H. I.; Pienta, K. J.; Kantoff, P.; de Bono, J. S.; Rubin, M. A.; Nelson, P. S.; Garraway, L. A.; Sawyers, C. L.; Chinnaiyan, A. M. Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell* **2015**, *161* (5), 1215−28.

(10) Schwarzenbach, H.; Alix-Panabieres, C.; Muller, I.; Letang, N.; Vendrell, J. P.; Rebillard, X.; Pantel, K. Cell-Free Tumor DNA in Blood Plasma as a Marker for Circulating Tumor Cells in Prostate Cancer. *Clin. Cancer Res.* **2009**, *15* (3), 1032−8.

(11) Lallous, N.; Volik, S. V.; Awrey, S.; Leblanc, E.; Tse, R.; Murillo, J.; Singh, K.; Azad, A. A.; Wyatt, A. W.; LeBihan, S.; Chi, K. N.; Gleave, M. E.; Rennie, P. S.; Collins, C. C.; Cherkasov, A. Functional Analysis of Androgen Receptor Mutations That Confer Anti-Androgen Resistance Identified in Circulating Cell-Free DNA from Prostate Cancer Patients. *Genome Biol.* **2016**, DOI: 10.1186/s13059-015-0864-1.

(12) Lavecchia, A. Machine-Learning Approaches in Drug Discovery: Methods and Applications. *Drug Discovery Today* **2015**, *20* (3), 318−31.

(13) Melagraki, G.; Afantitis, A. Editorial: Advances in Cheminformatics: Drug Discovery, Computational Toxicology and Nanomaterials (Part I). *Comb. Chem. High Throughput Screening* **2015**, *18* (3), 236−7.

(14) Lawless, M. S.; Waldman, M.; Fraczkiewicz, R.; Clark, R. D. Using Cheminformatics in Drug Discovery. *Handb. Exp. Pharmacol.* **2015**, *232*, 139−68.

(15) Mitchell, J. B. Machine Learning Methods in Chemoinformatics. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4* (5), 468−481.

(16) Yuriev, E.; Holien, J.; Ramsland, P. A. Improvements, Trends, and New Ideas in Molecular Docking: 2012−2013 in Review. *J. Mol. Recognit.* **2015**, *28* (10), 581−604.

(17) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, II; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. Qsar Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 4977−5010.

(18) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining Docking and Molecular Dynamic Simulations in Drug Design. *Med. Res. Rev.* **2006**, *26* (5), 531−68.

(19) Boyd, S. Molecular Operating Environment. *Chem. World* **2005**, *2* (9), 66−66.

(20) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. Integrated Modeling Program, Applied Chemical Theory (Impact). *J. Comput. Chem.* **2005**, *26* (16), 1752−80.

(21) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177−96.

(22) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739−49.

(23) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750−9.

(24) Clegg, N. J.; Wongvipat, J.; Joseph, J. D.; Tran, C.; Ouk, S.; Dilhas, A.; Chen, Y.; Grillot, K.; Bischoff, E. D.; Cai, L.; Aparicio, A.; Dorow, S.; Arora, V.; Shao, G.; Qian, J.; Zhao, H.; Yang, G.; Cao, C.; Sensintaffar, J.; Wasielewska, T.; Herbert, M. R.; Bonnefous, C.; Darimont, B.; Scher, H. I.; Smith-Jones, P.; Klang, M.; Smith, N. D.; De Stanchina, E.; Wu, N.; Ouerfelli, O.; Rix, P. J.; Heyman, R. A.; Jung, M. E.; Sawyers, C. L.; Hager, J. H. Arn-509: A Novel Antiandrogen for Prostate Cancer Treatment. *Cancer Res.* **2012**, *72* (6), 1494−503.

(25) Kennealey, G. T.; Furr, B. J. A. Use of the Nonsteroidal Antiandrogen Casodex in Advanced Prostatic-Carcinoma. *Urol. Clin. N. Am.* **1991**, *18* (1), 99−110.

(26) Neri, R.; Van Cleave, S.; Florance, K.; Koziol, P. Biological Profile of a Nonsteroidal Antiandrogen, Sch 13521 4′-Nitro-3′-Trifluoromethylisobutyranilide. *Endocrinology* **1972**, *91* (2), 427.

(27) Scher, H. I.; Fizazi, K.; Saad, F.; Taplin, M. E.; Sternberg, C. N.; Miller, K.; de Wit, R.; Mulders, P.; Chi, K. N.; Shore, N. D.; Armstrong, A. J.; Flaig, T. W.; Flechon, A.; Mainwaring, P.; Fleming, M.; Hainsworth, J. D.; Hirmand, M.; Selby, B.; Seely, L.; de Bono, J. S.; et al. Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. *N. Engl. J. Med.* **2012**, *367* (13), 1187−97.

(28) Todeschini, R.; Consonni, V.; Mannhold, R.; Kubinyi, H.; Folkers, G. Molecular Descriptors for Chemoinformatics. In *Methods and Principles in Medicinal Chemistry Series 82*; Wiley-VCH Imprint John Wiley & Sons, Incorporated: Hoboken, 2010; Vols. *I and II*, p 1.

(29) Dehmer, M.; Varmuza, K.; Bonchev, D. Statistical Modelling of Molecular Descriptors in Qsar/Qspr. In *Quantitative and network biology*; Wiley-Blackwell: Weinheim, 2012; Vol. *2*, p 1.

(30) Cherkasov, A. 'Inductive' Descriptors: 10 Successful Years in Qsar. *Curr. Comput.-Aided Drug Des.* **2005**, *1* (1), 21−42.

(31) Cherkasov, A. R.; Galkin, V. I.; Cherkasov, R. A. A New Approach to the Theoretical Estimation of Inductive Constants. *J. Phys. Org. Chem.* **1998**, *11* (7), 437−447.

(32) Cherkasov, A. R.; Galkin, V.; Cherkasov, R. "Inductive" Electronegativity Scale. *J. Mol. Struct.: THEOCHEM* **1999**, *489* (1), 43−46.

(33) Cherkasov, A. R.; Jonsson, M.; Galkin, V. A Novel Approach to the Analysis of Substituent Effects: Quantitative Description of Ionization Energies and Gas Basicity of Amines. *J. Mol. Graphics Modell.* **1999**, *17* (1), 28−42.

(34) Cherkasov, A.; Jonsson, M. A New Method for Estimation of Homolytic C-H Bond Dissociation Enthalpies. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (5), 1222−6.

(35) Cherkasov, A. R.; Galkin, V.; Cherkasov, R. "Inductive" Electronegativity Scale: 2. 'Inductive' Analog of Chemical Hardness. *J. Mol. Struct.: THEOCHEM* **2000**, *497*, 115−121.

(36) Andrews, C. G. *Scientific Vector Language as a Software Development Environment in Quantitative Structure Activity Relationships Studies*; University of New Brunswick, Faculty of Computer Science, 2003.

(37) *R: A Language and Environment for Stastitical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.

(38) *Rstudio: Integrated Development for R*; Boston, MA, 2015.

(39) Breiman, L. Random Forests. *Machine Learning* **2001**, *45* (1), 5−32.

(40) Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28* (5), 26.

(41) Kursa, M. B.; Jankowski, A.; Rudnicki, W. R. Boruta - a System for Feature Selection. *Fundamenta Informaticae* **2010**, *101* (4), 271−286.

(42) Verma, J.; Khedkar, V. M.; Coutinho, E. C. 3d-Qsar in Drug Design−a Review. *Curr. Top. Med. Chem.* **2010**, *10* (1), 95−115.

(43) Chang, C. C.; Lin, C. J. Libsvm: A Library for Support Vector Machines. *ACM TIST* **2011**, *2* (3), 1.

(44) Aha, D. W.; Kibler, D.; Albert, M. K. Instance-Based Learning Algorithms. *Machine Learning* **1991**, *6* (1), 37−66.

(45) Breiman, L. Bagging Predictors. *Machine Learning* **1996**, *24* (2), 123−140.

(46) Ting, K. M.; Witten, I. H. Stacking Bagged and Dagged Models. In *Proceedings of the Fourteenth International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., 1997; pp 367−375.

(47) Iba, W.; Langley, P. Induction of One-Level Decision Trees. *Machine Learning* **1992**, 233−240.

(48) Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I. H. The Weka Data Mining Software: An Update. *SIGKDD Explor. Newsl.* **2009**, *11* (1), 10−18.

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516

(49) Holte, R. C. Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Machine Learning* **1993**, *11* (1), 63−91.
(50) Sushko, I.; Novotarskyi, S.; Korner, R.; Pandey, A. K.; Cherkasov, A.; Li, J.; Gramatica, P.; Hansen, K.; Schroeter, T.; Muller, K. R.; Xi, L.; Liu, H.; Yao, X.; Oberg, T.; Hormozdiari, F.; Dao, P.; Sahinalp, C.; Todeschini, R.; Polishchuk, P.; Artemenko, A.; Kuz'min, V.; Martin, T. M.; Young, D. M.; Fourches, D.; Muratov, E.; Tropsha, A.; Baskin, I.; Horvath, D.; Marcou, G.; Muller, C.; Varnek, A.; Prokopenko, V. V.; Tetko, I. V. Applicability Domains for Classification Problems: Benchmarking of Distance to Models for Ames Mutagenicity Set. *J. Chem. Inf. Model.* **2010**, *50* (12), 2094−111.
(51) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3* (2), 198−210.
(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.*Gaussian 09*, Gaussian, Inc.: Wallingford, CT, USA, 2009.
(53) Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247−260.
(54) Wang, J.; Wang, W.; Kollmann, P.; Case, D. Antechamber, an Accessory Software Packagefor Molecular Mechanical Calculation. *J. Comput. Chem.* **2005**, *25*, 1157−1174.
(55) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.
(56) Harrach, M. F.; Drossel, B. Structure and Dynamics of Tip3p, Tip4p, and Tip5p Water near Smooth and Atomistic Walls of Different Hydroaffinity. *J. Chem. Phys.* **2014**, *140* (17), 174501.
(57) Steketee, K.; Timmerman, L.; Ziel-van der Made, A. C.; Doesburg, P.; Brinkmann, A. O.; Trapman, J. Broadened Ligand Responsiveness of Androgen Receptor Mutants Obtained by Random Amino Acid Substitution of H874 and Mutation Hot Spot T877 in Prostate Cancer. *Int. J. Cancer* **2002**, *100* (3), 309−17.
(58) Wahi, D.; Jamal, S.; Goyal, S.; Singh, A.; Jain, R.; Rana, P.; Grover, A. Cheminformatics Models Based on Machine Learning Approaches for Design of Usp1/Uaf1 Abrogators as Anticancer Agents. *Syst. Synth. Biol.* **2015**, *9* (1−2), 33−43.
(59) Derewenda, U.; Derewenda, Z.; Dodson, E. J.; Dodson, G. G.; Reynolds, C. D.; Smith, G. D.; Sparks, C.; Swenson, D. Phenol Stabilizes More Helix in a New Symmetrical Zinc Insulin Hexamer. *Nature* **1989**, *338* (6216), 594−6.

DOI: 10.1021/acs.jcim.6b00400
J. Chem. Inf. Model. 2016, 56, 2507−2516