2015

Modeling ATP-Binding Cassette G2 (ABCG2) Substrate Specificity

Raghav D. Acharya
Virginia Commonwealth University, acharyard@vcu.edu

Aurijit Sarkar
Virginia Commonwealth University

Glen E. Kellogg
Virginia Commonwealth University

VIRGINIA COMMONWEALTH UNIVERSITY

Modeling ATP-Binding Cassette G2 (ABCG2) Substrate Specificity

Raghav D. Acharya[1], Aurijit Sarkar[2], and Glen E. Kellogg[2*]
Department of Biology,[1] and Department of Medicinal Chemistry[2], VCU, Richmond, VA

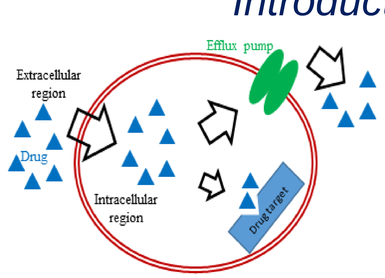
Abstract

- Cancer estimates for USA in 2015:
  - 1.6 million new cases,
  - half a million deaths [1]
  - majority of deaths due to resistance to chemotherapy [2]
- ATP-binding cassette (ABC) efflux transporters (e.g., ABCG2)
  - overexpressed in chemotherapy-resistant cancer cells
  - Anticancer drugs are prone to efflux
- What we need:
  - identify substrate and non-substrate chemotypes
  - gain a structural understanding of the efflux mechanism

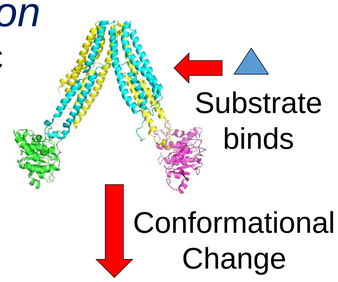
Aim: Understand ABCG2 structure and function

Introduction

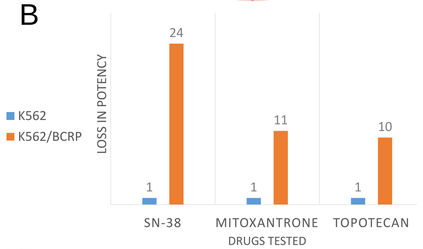
A

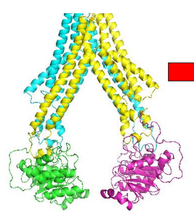


C



B

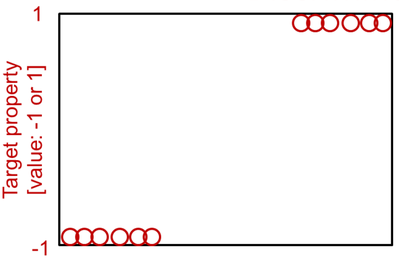




Why are certain compounds effluxed while others are not?

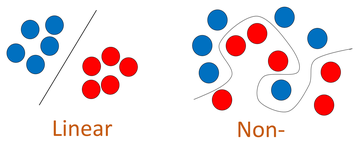
Method

Discrimination Analysis



Target property = ax + by + cz + k

Where:
a, b & c are correlation coefficients
x, y & z are independent properties



Non-Linear SVM method used for this study.

References

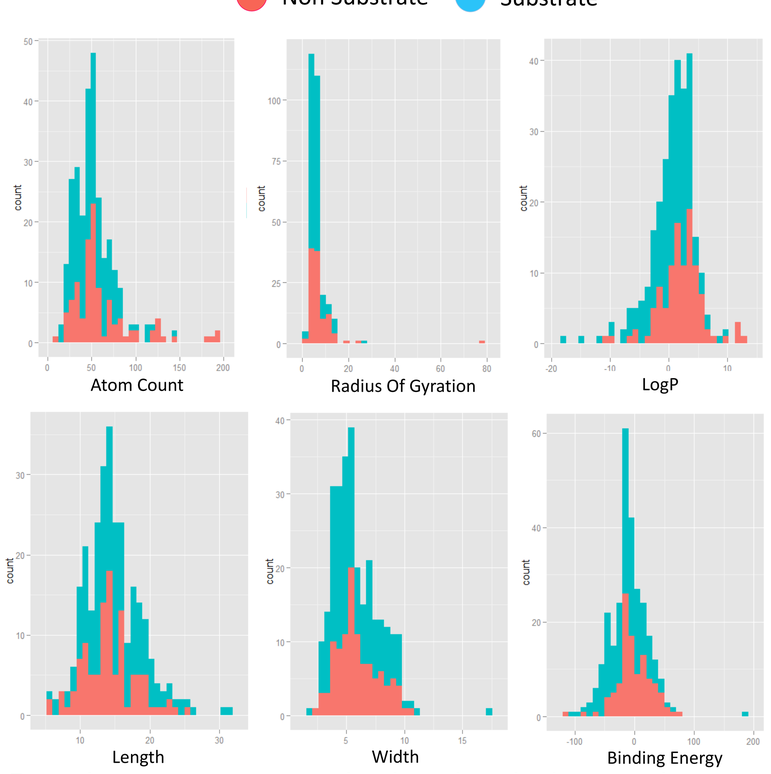
(1) American Cancer Society "facts and figures 2015" www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2015/index (accessed March 30, 2014).
(2) Gottesman, M. M. et al. *Nat. Rev. Cancer* **2002**, *2*, 48–58.
(3) Hazai, E. et al. *BMC Bioinformatics* **2013**, *14*, 130.
(B) Sugimoto et al. *Mol Cancer Ther* ***2003****, 2(1): 105-112*
(C) Rosenberg, M. F. et al. *Structure* ***2010***, *18*(4), 482–493.

Descriptors Used

Our Model
LogP (I)
Length (II)
Width (III)
Binding Energy (IV)
Atom Count (V)
Radius of Gyration (VI)

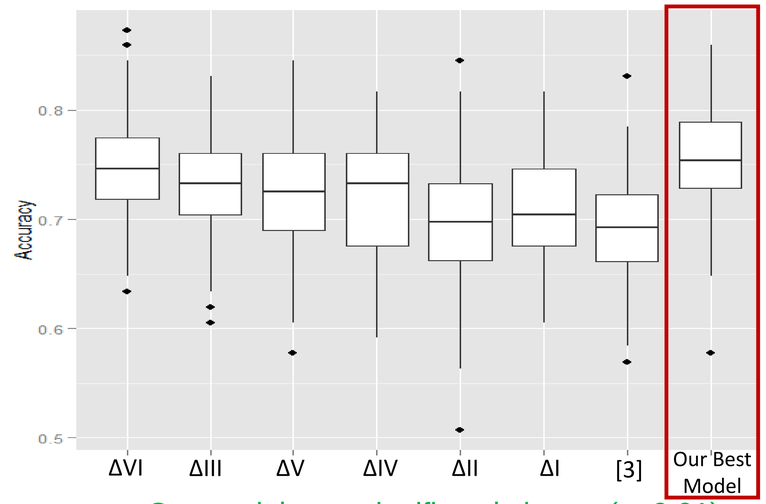
[3]
3D Morse signal 17/ weighed by mass
3D Morse signal 25/ weighed by mass
Gateway R autocorrelation of lag2 weighed by mass
Sperosity
Mean information on atomic composition

Lucid descriptors are capable of producing predictive models.



Descriptors encompass physico-chemical properties as well as efflux mechanism information.

Results



Our models are significantly better ($p<0.01$)

	Accuracy	fS	fNS
[3]	68.9 ± 4.9%	71.2 ± 7.1%	65.9 ± 14.8%
This work	***75.6 ± 4.7%***	***76.2 ± 6.5%***	***72.5 ± 11.1%***
Atom count	66.4 ± 5.2%	68.9 ± 7.1%	62.8 ± 17.1%
Binding energy	68.5 ± 5.4%	67.9 ± 5.8%	73.4 ± 13.9%
Radius Of Gyration	64.8 ± 4.7%	65.9 ± 5.9%	63.2 ± 19.3%
Length	65.3 ± 6.3%	67.7 ± 7.8%	50.2 ± 26.4%
Width	65.2 ± 5.4%	66.3 ± 6.5%	59.8 ± 23.1%
LogP	65.8 ± 6.1%	67.6 ± 7.2%	58.9 ± 26.3%

fS: fraction of substrates predicted correctly; ***fNS***: fraction of non-substrates predicted correctly; ***Accuracy***: fraction of dataset correctly predicted; Values shown in table represent Mean ± Standard Deviation of external validation set from 100 runs of SVM.

Discussion

- SVM model capable of discriminating between substrates and nonsubstrates with a median accuracy of 76.05% and an Interquartile range of 7.04%.
- Accuracy highly dependent on composition of training, test and external validation sets.
- Insights into efflux mechanism – role of Arg482 in substrate recognition suggested by significant difference in binding energy between substrates and non substrates.

Implications

- Discriminant models are noisy –understanding of the structural mechanism of efflux might lead to better models.
- More experimental data needed – might make for a better predictive model.

Future directions

- Glean structural information on ABCG2-mediated efflux to improve model.

Acknowledgements

This work was funded by the Mary Louise Andrews Award from the Virginia Academy of Science.