ISSN No:-2456-2165

Carbonic Anhydrase QSAR Model as Templates for Biomarkers Discovery

Lungu N. Claudiu[1]
Department of Chemistry
Faculty of Chemistry and Chemical Engineering
Babes-Bolyai University, Romania. 400028Cluj

Abstract:- Carbonic anhydrase (CA) has a significant role in abnormal cellular proliferation and might serve as a novel biomarker for cancer. Cellular level studies of several CA isoforms for CA inhibition potential had shown the role of CA in cancer pathways. The distinct tumor-associated CA isoforms identified in humans can be the targets for the potential approach in cancer therapeutics. Targeting these distinct isoforms in human tumorigenesis, by identifying CA inhibitors from plants, can be the answer to an advanced therapeutic strategy through personal therapy. Novel biomarkers of CA isoforms based on gene data and experimental results retrieved from literature to model a 3D protein sequence were developed using computational strategies.

Keywords:- *Carbonic Anhydrase, Biomarker, QSAR.*

I. INTRODUCTION

Carbonic anhydrase ia a well studied enzyme. The enzyme is targeted by some drugs, such as acetazolamide, methazolamide. There enzyme families were identified: alpha, beta, and gamma. All study CAs from the animal kingdom is of the alpha type.Fourteen different carbonic anhydrase isoforms were isolated in higher vertebrates.[1] Several essential functions are played by many CA isozymes, which are strongly constrained by aromatic and heterocyclic sulfonamides along with as inorganic, metal complexing anions[3]. The mode of action of these enzymes is acquainted in detail, and this helped the design of potent molecules that block their activity. A current finding is related to the engagement of CAs and their sulfonamide inhibitors in cancer. Many potent sulfonamide inhibitors inhibited the growth of a multitude of tumor cells in vitro and in vivo. A few types of compounds that form protein ligan complexes with CAs have been identified recently, those compounds possess modified sulfonamide or hydroxamate functional groups. In this computational study based on secondary derivate data, an activity model for Ca is refined. This model is to be used in further biomarkers discovery and design.

II. METHODS

In order to develop biomarker templates based on carbonic anhydrase, human CA was computationally analyzed. PDB structures for human CA from I to XIV were considered (1azm[3], 12ca, 1z93[4], 1znc, 1dmx, 3fe4[5], 3mdz[6], 2w2j[7], 5fl4[8], 1jcz,4knn[9], 5cjf[10]).Ligands, solvent molecules, and dummy atoms were removed from all discussed structures. Only subunit A was retained in all enzymes. Using the AMBER 99 force field, the energy was minimized, charges corrected, the structure was protonated at pH=7.4and 310K, and a salt concentration (NaCl) of 0.1mol/L, respectively. Human CA PDB structure for CA X, XI, XIII are not yet available.

A homology model having CA VII as a template (PDB id 3ml5) was generated for CA X, starting from its amino acid sequence. Sequence identity was 31.78%, GMQE 0.59, QMEAN -2.23. CA XI was modeled using CA II as the template (PDB id 4pxx). Sequence identity was 33.9%, GMQE 0.58, QMEAN -2.30.

In order to characterize biomarkers in terms of specific activity, the future biomarkers design, based on CA isoforms, a prediction model was developed. The model was built using multiple linear regression. The specific activity of human CA (s-1) was set as the dependent variable. Specific activity for each CA isoform (I-IXIV) was retrieved from the literature[17,18]. For all CA molecules, monomer form was used.

Descriptors used for each CA (I-XIV), were computed using Schrodinger[11], MOE[12], and Mathematica[13] software. In computing these descriptors, each CA isoform was converted into a molecular graph using Hyper-Chem[14] and TopoCluj[15]software.

QSAR model for predicting CA bioactivity was built using multiple linear regression (MLR). Validation of the model was performed using Leave one out technique (LOO). The 14 molecules were spread into a training and a test set respectively.
Descriptors cluster was represented in 2D plots(Figs. 1 and 2), for testing the descriptor interrelation.

To further explore the descriptor space, polynomial order, six equations were built using bond length and dihedral angle for each protein. Using these equations, areas

ISSN No:-2456-2165

under the curve were computed. Those equations were used to correlate with descriptors used in the QSAR model.

III. RESULTS

3D models for CA (carbonic anhydrase) 1-14 are shown in table 1. In figures 1 and 2 regression equation of the QSAR model and descriptors space is shown.

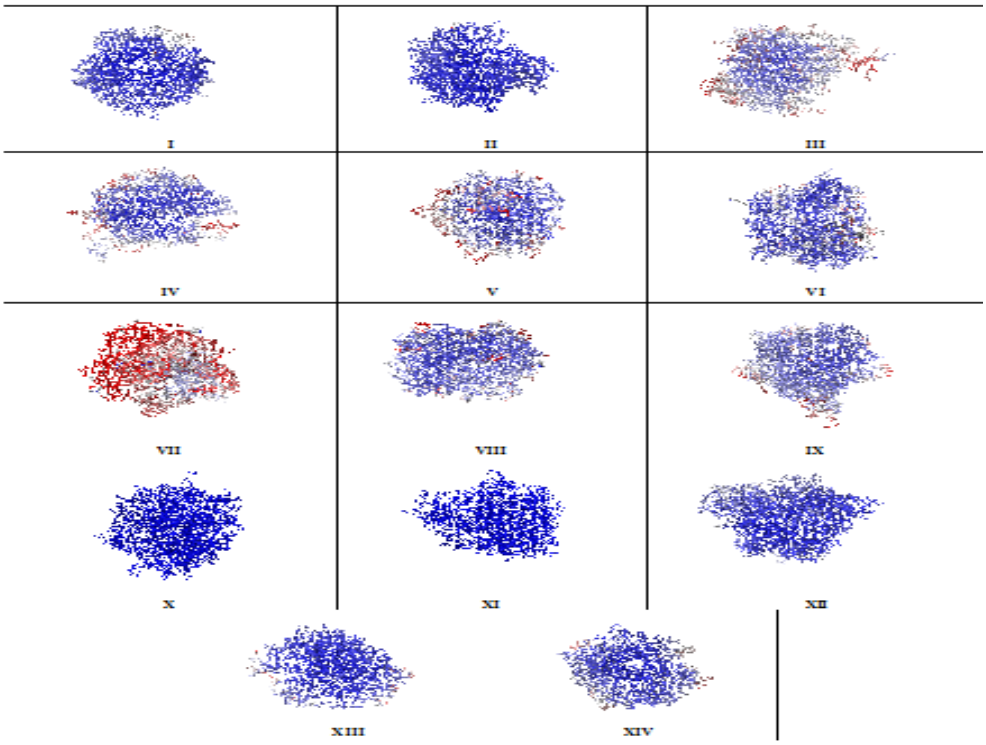

Table 1:- Carbonic Anhydrase Isoforms - B Factor Represented as Balls and Sticks

Model was built using multiple linear regression. Equation and descriptors used are presented in Table 2.

N	r^2	*Descriptors*
1	***0.99***	***BCUTSMR1, PEOEVSA+6, SlogPVSA3, chiral, vsurfDW12, vsurfEDmin3, vsurfIW8***
2	***0.98***	***BCUTSMR1, PEOEVSA+6, SlogPVSA3, chiral, vsurfDW12, vsurfEDmin3,***
3	***0.97***	***BCUTSMR1, PEOEVSA+6, chiral, vsurfDW12, vsurfEDmin3, vsurfIW8,***
4	***0.96***	***BCUTSMR1, PEOEVSA+6, chiral, vsurfDW12, vsurfEDmin3,***
5	***0.95***	***BCUTSMR1, GCUTPEOE0, PEOEVSA+6, chiral, vsurf_DW12,***
4	***0.94***	***BCUTSMR1, PEOEVSA+6, chiral, rsynth, vsurfDW12,***
5	***0.92***	***BCUTSMR1, PEOEVSA+6, chiral, vsurfDW12,***
6	***0.91***	***BCUTSMR1, PEOEVSA+6, rsynth, vsurf_DW12,***
7	***0.88***	***BCUTSMR1, PEOEVSA+6, vsurf_DW12,***
8	***0.83***	***BCUTSMR1, rsynth, vsurfDW12,***
9	***0.76***	***BCUTSMR1, vsurfDW12,***
10	***0.57***	***vsurfDW12,***

Table 2:- Multiple linear regression model built with the following descriptors: BCUTSMR1-The BCUT descriptors with atomic contribution to molar refractivity instead of partial charge, PEOEVSA+6 -Sum of vi where qi is greater than 0.3, SlogPVSA3-Sum of vi such that Li is in (0,0.1], chiral-The number of chiral centers, vsurf_DW12-Contact distances of vsurfEWmin3, vsurfEDmin3-Lowest hydrophobic energy, vsurfIW8-Hydrophilic integy moment (8 descriptors), rsynth- synthetic fesabiliy descriptors.

ISSN No:-2456-2165

Model resulted from 7 descriptors (BCUT_SMR_1, PEOE_VSA+6, SlogP_VSA3, chiral, vsurf_DW12, vsurf_EDmin3, vsurf_IW8) has r^2= 0.99, p-0.98, MSD-1.29, q2= 0.99 and the prediction equation: y=3888.3+0.993042×Observed activity.

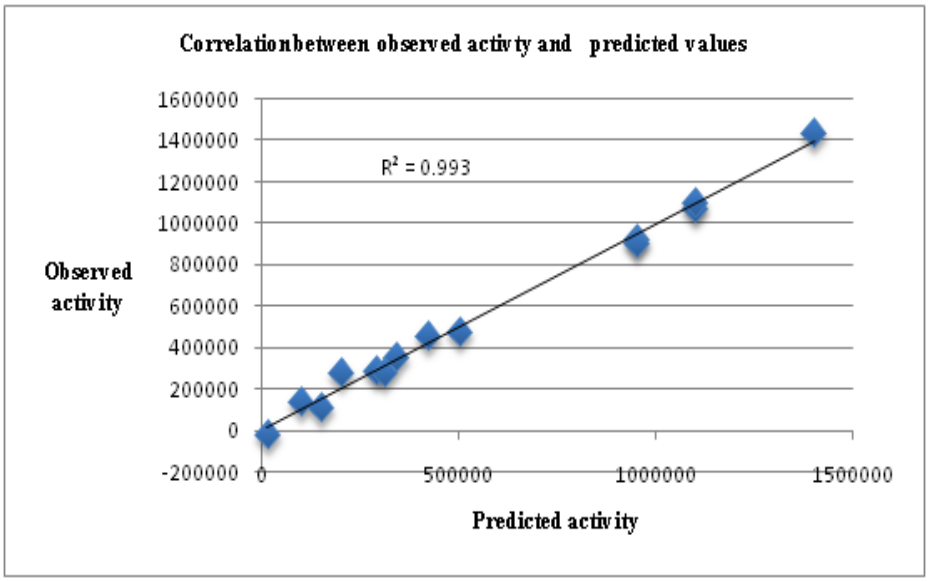

Fig 1:- QSAR Model for Carbonic Anhydrase

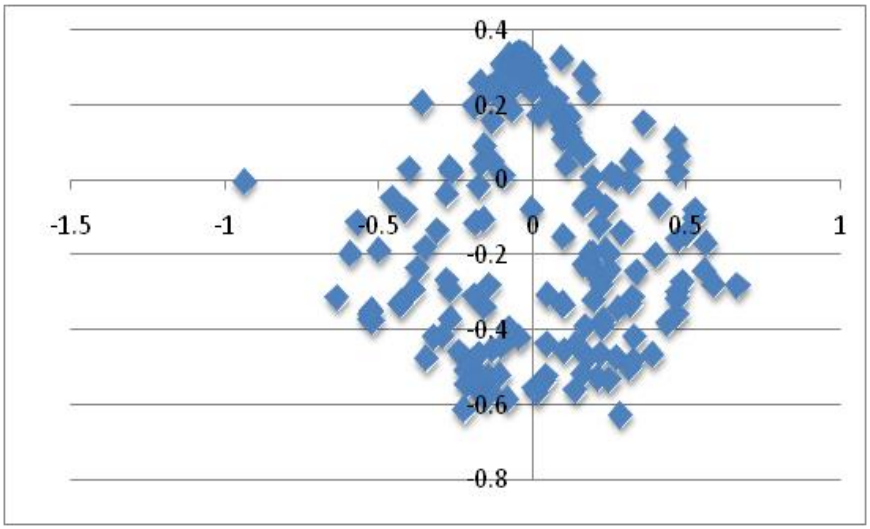

Fig 2:- Cluster of QSAR Descriptors

Model	*Bond length polynomial equation*	*Dihedral angle equation*
CA 1	$y = -19x^6 -14x^5 -11x^4 -0.8x^3 -0.5x^2 - 0.0197x + 2.6441$	$y = -18x^6 -14x^5 -11x^4 -0.7x^3 + 0.0003x^2 - 0.1334x + 12.979$
CA 2	$y = -18x^6 -14x^5 -11x^4 -0.9x^3 -05x^2 - 0.0088x + 2.2213$	$y = -18x^6 + 14x^5 -11x^4 -0.7x^3 - 0.0002x^2 + 0.0695x - 8.0386$
CA 3	$y = -18x^6 -14x^5 + 11x^4 -0.7x^3 + 0.0002x^2 - 0.0605x + 6.2892$	$y = -19x^6 -15x^5 -12x^4 -0.8x^3 -0.6x^2 + 0.0053x - 10.015$
CA 4	$y = -19x^6 -14x^5 -11x^4 -07x^3 -0.4x^2 - 0.0355x + 4.4062$	$y = -19x^6 -15x^5 -11x^4 -0.7x^3 + 0.0002x^2 - 0.0668x - 8.5229$
CA 5	$y = -18x^6 -14x^5 -11x^4 -0.8x^3 + 0.5x^2 - 0.0038x + 1.586$	$y = -19x^6 -15x^5 -11x^4 -0.8x^3 + 0.0002x^2 - 0.1765x + 51.341$
CA 6	$y = -20x^6 -16x^5 -12x^4 -0.9x^3 -0.7x^2 + 0.0001x + 1.3189$	$y = -18x^6 -14x^5 -11x4 -0.7x^3 + 0.0003x^2 - 0.1701x + 26.197$

CA 7	$y = -19x^6 -14x^5 -11x^4 -0.7x^3 -05x^2 - 0.0276x + 3.4486$	$y = -19x^6 -15x^5 -11x^4 -0.8x^3 -0.5x^2 + 0.0126x + 8.5378$
CA 8	$y = -17x^6 -13x^5 -10x^4 -0.7x^3 - 0.0003x^2 + 0.0737x - 2.372$	$y = -19x^6 -15x^5 -11x^4 -0.8x^3 -0.5x^2 - 0.0334x + 16.88$
CA 9	$y = -20x^6 -15x^5 -11x^4 -0.8x^3 -0.5x^2 + 0.0136x + 0.2563$	$y = -17x^6 -13x^5 -10x^4 -0.7x^3 - 0.0003x^2 + 0.0179x + 28.311$
CA 10	$y = -19x^6 -15x^5 -12x^4 -0.9x^3 -0.6x^2 - 0.0005x + 1.4446$	$y = -18x^6 -14x^5 -11x^4 -0.7x^3 + 0.0002x^2 - 0.12x + 30.724$
CA 11	$y = -18x^6 -15x^5 -11x^4 -0.8x^3 -0.6x^2 - 0.0013x + 1.4871$	$y = -18x^6 -14x^5 -11x^4 -0.7x^3 + 0.0002x^2 - 0.109x + 31.855$
CA 12	$y = -19x^6 -15x^5 -11x^4 -0.8x^3 -0.5x^2 - 0.004x + 1.1516$	$y = -19x^6 -15x^5 -13x^4 -0.8x^3 -0.5x^2 - 0.0583x + 3.1948$
CA 13	$y = -18x^6 -14x^5 -11x^4 -0.8x^3 -0.5x^2 - 0.0029x + 1.5573$	$y = -19x^6 -15x^5 -11x^4 -0.7x^3 + 0.0002x^2 - 0.0984x + 29.467$
CA 14	$y = -19x^6 -15x^5 -11x^4 -08x^3 -0.5x^2 - 0.0269x + 3.4656$	$y = -19x^6 -15x^5 -11x^4 -0.8x^3 + 0.0001x^2 - 0.0689x + 11.984$

Table 3:- Polynomial Equations for CA Monomers

For each CA monomer a distinct equation was obtained -Table 3). Areas under the curve resulted based on each equation are shown in table 4.

Model	***Bond length polynomial equation based area***	***Dihedral angle equation-based area***
CA 1	***2.71746***	***19.1986***
CA 2	***1.65166***	-
CA 3	***15.7216***	-
CA 4	***5.70645***	-
CA 5	***1.58954***	***98.3808***
CA 6	***1.11471***	***44.3944***
CA 7	***2.99641***	***11.4111***
CA 8	-	***25.9406***
CA 9	***0.143058***	***49.3381***
CA 10	***1.25969***	***53.656***
CA 11	***1.33504***	***56.0112***
CA 12	***0.974505***	***3.34435***
CA 13	***1.41803***	***50.8769***
CA 14	***4.52706***	***17.4196***

Table 4:- Areas Based on Polynomial Equations ($Å^2$)

QSAR model build using MLR shows that enzyme bioactivity correlates with ***BCUTSMR1, PEOEVSA+6, chiral, rsynth, vsurf_DW12.*** descriptors. Enzyme activity is related to surface and shape. ***chiral, rsynth*** probably are showing the same impact on the shape and molecular surface regarding bioactivity.

Undetermined surfaces are due to unbounded functions. In calculus, a function f describe on any set X with complex values is named bounded if the collection of its values is bounded. Namely, there must be a real number M such that $f(x) \leq M$ for entirely x in X[16]. If f(x) is not bounded is said to be unbounded[17].

The surface under the curve computed with bond lengths polynomial equation is related to ***GCUT,*** and ***Slog*** descriptors and surface under the curve computed based on dihedral angles equation also relates to CGUT.

QSAR model will be used to further predict bioactivity for peptide motifs constructed using equations obtained for bond length and dihedral angle used in this study in order to develop and identify the future biomarkers. Peptide motifs will be further developed using bond lengths and dihedral angles equations.

IV. CONCLUSIONS

MLR model predicted accurately CA bioactivities, retrieved from literature. Bioactivity of CA isoforms is shape and surface dependent. The model will be further used to predict bioactive of novel enhance bioactivity - peptide motifs.

REFERENCES

[1]. Hilvo M, Tolvanen M, Clark A, Shen B, Shah GN, Waheed A, Halmi P, Hänninen M, Hämäläinen JM, Vihinen M, Sly WS, Parkkila S. Characterization of CA XV, a new GPI-anchored form of carbonic anhydrase. Biochem. J. 2005;392 83–92.

[2]. Briganti, Fabrizio & Tilli, Silvia & Mincione, Giovanna & Mincione, Francesco & Menabuoni, Luca & Supuran, Claudiu. (2000). Carbonic Anhydrase Inhibitors. Metal Complexes of 5-(2-Chlorophenyl)-1, 3, 4-Thiadiazole-2-Sulfonamide with Topical Intraocular Pressure Lowering Properties: The Influence of Metal Ions Upon the Pharmacological Activity. Journal of enzyme inhibition. 2000;15: 185-200.

[3]. Chakravarty, S., Kannan, K.K. Drug-protein interactions.Refined structures of three sulfonamide drug complexes of human carbonic anhydrase I enzyme.J.Mol.Biol. 1994;243: 298-309.

[4]. Duda, D.M., Tu, C., Fisher, S.Z., An, H., Yoshioka, C., Govindasamy, L., Laipis, P.J., Agbandje-McKenna, M., Silverman, D.N., McKenna, R.. Human Carbonic Anhydrase III: Structural and Kinetic Study of Catalysis and Proton TransferBiochemistry 2005;44: 10046-10053.

[5]. Pilka, E.S., Kochan, G., Oppermann, U., Yue, W.W. Crystal structure of the secretory isozyme of mammalian carbonic anhydrases CA VI: implications for biological assembly and inhibitor development.Biochem.Biophys.Res.Commun. 2012;419: 485-489.

[6]. Ugochukwu, E., Shafqat, N., Pilka, E., Chaikuad, A., Krojer, T., Muniz, J., Kim, J., Bray, J., Bountra, C., Arrowsmith, C.H., Weigelt, J., Edwards, A., von Delft, F., Carpenter, E.P., Yue, W.W., Oppermann,, Crystal Structure of Human Carbonic Anhydrase VII [isoform 1], CA7, to be published (2018)

[7]. Picaud, S.S., Muniz, J.R.C., Kramm, A., Pilka, E.S., Kochan, G., Oppermann, U., Yue, W.W. Crystal Structure of Human Carbonic Anhydrase-Related Protein Viii Reveals the Basis for Catalytic Silencing.Proteins 2009;76: 507.

[8]. Leitans, J., Kazaks, A., Balode, A., Ivanova, J., Zalubovskis, R., Supuran, C.T., Tars, K.An Efficient Expression and Crystallization System of the Cancer-Associated Carbonic Anhydrase Isoform Ix.J.Med.Chem. 2015;58: 9004.

[9]. Capkauskaite, E., Zubriene, A., Smirnov, A., Torresan, J., Kisonaite, M., Kazokaite, J., Gylyte, J., Michailoviene, V., Jogaite, V., Manakova, E., Grazulis, S., Tumkevicius, S., Matulis, D.Benzenesulfonamides with pyrimidine moiety as inhibitors of human carbonic anhydrases I, II, VI, VII, XII, and XIII, Bioorg.Med.Chem. 2013;21: 6937-6947.

[10]. La Regina, G., Coluccia, A., Famiglini, V., Pelliccia, S., Monti, L., Vullo, D., Nuti, E., Alterio, V., De Simone, G., Monti, S.M., Pan, P., Parkkila, S., Supuran, C.T., Rossello, A., Silvestri, R., Discovery of 1,1'-Biphenyl-4-sulfonamides as a New Class of Potent and Selective Carbonic Anhydrase XIV Inhibitors.J.Med.Chem. 2015;58: 8564-8572.

[11]. Hilvo M, Baranauskiene L, Salzano AM, Scaloni A, Matulis D, Innocenti A, Scozzafava A, Monti SM, Di Fiore A, De Simone G, Lindfors M, Jänis J, Valjakka J, Pastoreková S, Pastorek J, Kulomaa MS, Nordlund HR, Supuran CT, Parkkila S (2008).Biochemical characterization of CA IX, one of the most active carbonic anhydrase isozymes. J. Biol. Chem. 2008;283 (41): 27799–809.

[12]. SilviaPastorekova, SeppoParkkila, Jan Zavada. Tumor-associated Carbonic Anhydrases and Their Clinical Significance.Advances in Clinical Chemistry. 2006;42:167-216.

[13]. Sherman, W. Day, T., Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. Journal of Medicinal Chemistry. 2006;49 (2): 534–553.

[14]. Erlanson DA, McDowell RS, O'Brien T.Fragment-based drug discovery. J. Med. Chem. 2004;47 (14): 3463–82.

[15]. Wolfram Research, Inc., Mathematica, Version 11.3, Champaign, IL, 2018.

[16]. William F. Coleman, and Christopher R. Arumainayagam.HyperChem 5 (by Hypercube, Inc.)J. Chem. Educ., 1998;75 (4): 416.

[17]. M., V.Diudea, A. E. Vizitiu.Omega polynomial in twisted (4,4) tori.Match-Communications Mathematical and in Computer Chemistry, 2008;60(3): 945-953.